



```

; LENGTH: 103
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-489-448-1704

Alignment Scores:
Pred. No.: 1.88 Length: 103
Score: 54.50 Matches: 8
Percent Similarity: 61.54% Conservative: 0
Best Local Similarity: 61.54% Mismatches: 4
Query Match: 43.25% Indels: 1
DB: 6 Gaps: 1

US-10-775-481A-4 (1-57) x US-10-489-448-1704 (1-103)
QY 16 TGCTGCTGAATTGTTGTTGTAATCTGCTGTTGTAACGGGTGC 54
Db 45 CysCysSerCysCysCys---ProCysCysArgGlyCys 56

RESULT 7
US-10-938-061-106
; Sequence 106, Application US/10938061
; GENERAL INFORMATION:
; APPLICANT: Cairns, Belinda
; APPLICANT: Chen, Ruihuan
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Koeppeen, Hartmut
; APPLICANT: Phillips, Heidi S.
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan D.
; APPLICANT: Smith, Victoria
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas D.
; APPLICANT: Zhang, Zemin
; APPLICANT: Chantharapai, Anan
; APPLICANT: Reed Chae J.
; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; TITLE OF INVENTION: Treatment of Tumor
; FILE REFERENCE: P5001R1PIB
; CURRENT APPLICATION NUMBER: US/10/938,061
; PRIOR FILING DATE: 2004-09-10
; PRIOR APPLICATION NUMBER: US 10/872,991
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/872,972
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/241,220
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 60/299,500
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/301,880
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/323,268
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US 60/557,116
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: US 60/598,899
; PRIOR FILING DATE: 2004-08-04
; NUMBER OF SEQ ID NOS: 154
; SEQ ID NO 106
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-938-061-106

Alignment Scores:
Pred. No.: 2.22 Length: 690
Score: 54.00 Matches: 7
Percent Similarity: 57.14% Conservative: 1
Best Local Similarity: 57.14% Mismatches: 6
Query Match: 50.00% Indels: 0
DB: 6 Gaps: 0

US-10-775-481A-4 (1-57) x US-10-938-061-106 (1-690)
QY 13 TACTGCTGTGAATTGTTGTTGTAATCTGCTGTTGTAACGGGTGC 54
Db 620 TyrCysCysArgValCysArgAlaCysCysLeuLeuCys 633

RESULT 9
PCT-US05-05406-6
; Sequence 6, Application PC/TUS0505406
; GENERAL INFORMATION:
; APPLICANT: Cairns, Belinda
; APPLICANT: Chen, Ruihuan
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Koeppeen, Hartmut
; APPLICANT: Phillips, Heidi S.
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan D.
; APPLICANT: Smith, Victoria
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas D.
; APPLICANT: Zhang, Zemin
; APPLICANT: Chantharapai, Anan
; APPLICANT: Reed Chae J.
; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; TITLE OF INVENTION: Treatment of Tumor
; FILE REFERENCE: P5001R1PIB
; CURRENT APPLICATION NUMBER: US/10/936,626
; PRIOR FILING DATE: 2004-09-08
; PRIOR APPLICATION NUMBER: US 10/872,991
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/872,972
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/241,220
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 10/177,488
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US 60/299,500
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/301,880
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/323,268
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US 60/557,116
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: US 60/598,899
; PRIOR FILING DATE: 2004-08-04
; NUMBER OF SEQ ID NOS: 154
; SEQ ID NO 106
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-936-626-106

Alignment Scores:
Pred. No.: 2.22 Length: 690
Score: 54.00 Matches: 7
Percent Similarity: 57.14% Conservative: 1
Best Local Similarity: 50.00% Mismatches: 6
Query Match: 42.86% Indels: 0
DB: 6 Gaps: 0

US-10-775-481A-4 (1-57) x US-10-936-626-106 (1-690)
QY 13 TACTGCTGTGAATTGTTGTTGTAATCTGCTGTTGTAACGGGTGC 54
Db 620 TyrCysCysArgValCysArgAlaCysCysLeuLeuCys 633

RESULT 9
PCT-US05-05406-6
; Sequence 6, Application PC/TUS0505406
; GENERAL INFORMATION:
```

```
; APPLICANT: Olsen, Byron
; TITLE OF INVENTION: A Method for the Production of Transgenic Proteins Useful in the
; FILE REFERENCE: GTC-223 PCT
; CURRENT APPLICATION NUMBER: PCT/US05/05406
; PRIOR FILING DATE: 2005-02-18
; PRIOR APPLICATION NUMBER: 60/545,790
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 1071
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-05406-6

Alignment Scores:
Pred. No.: 2.22 Length: 1071
Score: 54.00 Matches: 7
Percent Similarity: 53.85% Conservative: 0
Best Local Similarity: 53.85% Mismatches: 6
Query Match: 42.86% Indels: 0
DB: 1 Gaps: 0

US-10-775-481A-4 (1-57) x PCT-US05-05406-6 (1-1071)
Qy 16 TGCTGTGAATGTGTTGTAATCCTGCTTGTAAACGGGTGC 54
Db 261 CysCysGlyAlaCysCysThrCysThrCysGlyGlyCys 273

RESULT 10
US-11-060-291-6
; Sequence 6, Application US/11060291
; GENERAL INFORMATION:
; APPLICANT: Olsen, Byron
; TITLE OF INVENTION: A Method for the Production of Transgenic Proteins Useful in the
; FILE REFERENCE: GTC-223
; CURRENT APPLICATION NUMBER: US/11/060,291
; PRIOR FILING DATE: 2005-02-17
; PRIOR APPLICATION NUMBER: 60/545,790
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 1071
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-060-291-6

Alignment Scores:
Pred. No.: 2.22 Length: 1071
Score: 54.00 Matches: 7
Percent Similarity: 53.85% Conservative: 0
Best Local Similarity: 53.85% Mismatches: 6
Query Match: 42.86% Indels: 0
DB: 7 Gaps: 0

US-10-775-481A-4 (1-57) x US-11-060-291-6 (1-1071)
Qy 16 TGCTGTGAATGTGTTGTAATCCTGCTTGTAAACGGGTGC 54
Db 261 CysCysGlyAlaCysCysThrCysThrCysGlyGlyCys 273

RESULT 11
US-10-287-436A-285
; Sequence 285, Application US/10287436A
; GENERAL INFORMATION:
; APPLICANT: CHILDREN'S HOSPITAL MEDICAL CENTER
; TITLE OF INVENTION: METHOD FOR DIAGNOSIS AND TREATMENT OF
; FILE REFERENCE: 10872.514696
; CURRENT APPLICATION NUMBER: US/10/287.436A
```

```
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/336,220
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 1446
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 285
; LENGTH: 1123
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-287-436A-285

Alignment Scores:
Pred. No.: 2.22 Length: 1123
Score: 54.00 Matches: 7
Percent Similarity: 53.85% Conservative: 0
Best Local Similarity: 53.85% Mismatches: 6
Query Match: 42.86% Indels: 0
DB: 6 Gaps: 0

US-10-775-481A-4 (1-57) x US-10-287-436A-285 (1-1123)
Qy 16 TGCTGTGAATGTGTTGTAATCCTGCTTGTAAACGGGTGC 54
Db 64 CysCysAlaGlyCysCysThrCysThrCysGlyCys 76

RESULT 12
PCT-US05-03560-6
; Sequence 6, Application PC/TUS0503560
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: ITPKS AS SMODIFIERS OF THE IGFR PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX05-004C-PC
; CURRENT APPLICATION NUMBER: PCT/US05/03560
; CURRENT FILING DATE: 2005-01-27
; PRIOR APPLICATION NUMBER: US60/539,837
; PRIOR FILING DATE: 2004-01-28
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 3398
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-03560-6

Alignment Scores:
Pred. No.: 2.2 Length: 3398
Score: 54.00 Matches: 7
Percent Similarity: 53.85% Conservative: 0
Best Local Similarity: 53.85% Mismatches: 6
Query Match: 42.86% Indels: 0
DB: 1 Gaps: 0

US-10-775-481A-4 (1-57) x PCT-US05-03560-6 (1-3398)
Qy 16 TGCTGTGAATGTGTTGTAATCCTGCTTGTAAACGGGTGC 54
Db 2463 CysCysThrCysCysCysThrGlyThrCysAlaGlyCys 2475

RESULT 13
US-60-643-717-5307
; Sequence 5307, Application US/60643717
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)A
; CURRENT APPLICATION NUMBER: US/60/643,717
; CURRENT FILING DATE: 2005-01-12
; NUMBER OF SEQ ID NOS: 19247
; SEQ ID NO 5307
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Kluyveromyces lactis
US-60-643-717-5307
```

Alignment Scores:  
Pred. No.: 3.16 Length: 423  
Score: 53.00 Matches: 10  
Percent Similarity: 66.67% Conservative: 0  
Best Local Similarity: 66.67% Mismatches: 5  
Query Match: 50.96% Indels: 0  
DB: 8 Gaps: 0

US-10-775-481A-4 (1-57) x US-60-643-717-5307 (1-423)

QY 46 TACAAGCAGGATTACACACATTACAGCAGTAACTGCTACTAT 2  
|||||  
Db 77 TyrlYsGlnSerLysLeuIleHisSerSerAsnLeuTyTyTyr 91  
|||||

## RESULT 14

US-60-655-875-169174  
; Sequence 169174, Application US/60655875  
; GENERAL INFORMATION:  
; APPLICANT: Boukharov, Andrey  
; APPLICANT: Du, Zijiang  
; APPLICANT: Guo, Liang  
; APPLICANT: Kovalic, David  
; APPLICANT: Lu, Maolong  
; APPLICANT: McCarter, James  
; APPLICANT: Miller, Nancy  
; APPLICANT: Williams, Deryck  
; APPLICANT: Vaudin, Mark  
; APPLICANT: Wu, Wei  
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS  
; FILE REFERENCE: 38-21(53885)  
; CURRENT APPLICATION NUMBER: US/60/655,875  
; CURRENT FILING DATE: 2005-02-24  
; NUMBER OF SEQ ID NOS: 171306  
; SEQ ID NO 169174  
; LENGTH: 152  
; TYPE: PRT  
; ORGANISM: Heterodera glycines  
; FEATURE:  
; OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID\_95597; Strand=-; Position=3-  
US-60-655-875-169174

Alignment Scores:  
Pred. No.: 4.52 Length: 152  
Score: 52.00 Matches: 7  
Percent Similarity: 61.54% Conservative: 1  
Best Local Similarity: 53.85% Mismatches: 5  
Query Match: 41.27% Indels: 0  
DB: 8 Gaps: 0

US-10-775-481A-4 (1-57) x US-60-655-875-169174 (1-152)

QY 10 AATTACTGCTGTAATTGTTGTTGTAATCCTGCTGTAAC 48  
|||||  
Db 11 AsnIleCysCysArgArgCysCysAsnCysCysCysSer 23  
|||||

## RESULT 15

US-60-655-875-163386  
; Sequence 163386, Application US/60655875  
; GENERAL INFORMATION:  
; APPLICANT: Boukharov, Andrey  
; APPLICANT: Du, Zijiang  
; APPLICANT: Guo, Liang  
; APPLICANT: Kovalic, David  
; APPLICANT: Lu, Maolong  
; APPLICANT: McCarter, James  
; APPLICANT: Miller, Nancy  
; APPLICANT: Williams, Deryck  
; APPLICANT: Vaudin, Mark  
; APPLICANT: Wu, Wei  
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS  
; FILE REFERENCE: 38-21(53885)  
; CURRENT APPLICATION NUMBER: US/60/655,875  
; CURRENT FILING DATE: 2005-02-24  
; NUMBER OF SEQ ID NOS: 171306  
; SEQ ID NO 163386  
; LENGTH: 158  
; TYPE: PRT  
; ORGANISM: Heterodera glycines  
; FEATURE:  
; OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID\_89809; Strand=-; Position=0-12  
US-60-655-875-163386

FILE REFERENCE: 38-21(53885)  
; CURRENT APPLICATION NUMBER: US/60/655,875  
; CURRENT FILING DATE: 2005-02-24  
; NUMBER OF SEQ ID NOS: 171306  
; SEQ ID NO 163386  
; LENGTH: 158  
; TYPE: PRT  
; ORGANISM: Heterodera glycines  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (2)..(2)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (4)..(4)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (14)..(14)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (18)..(21)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
; FEATURE:  
; OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID\_89809; Strand=-; Position=0-12  
US-60-655-875-163386

Alignment Scores:  
Pred. No.: 4.52 Length: 158  
Score: 52.00 Matches: 7  
Percent Similarity: 61.54% Conservative: 1  
Best Local Similarity: 53.85% Mismatches: 5  
Query Match: 41.27% Indels: 0  
DB: 8 Gaps: 0

US-10-775-481A-4 (1-57) x US-60-655-875-163386 (1-158)

QY 10 AATTACTGCTGTAATTGTTGTTGTAATCCTGCTGTAAC 48  
|||||  
Db 41 AsnIleCysCysArgArgCysCysAsnCysCysCysSer 53  
|||||

Search completed: March 26, 2005, 18:03:12  
Job time : 30 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model  
Run on: March 26, 2005, 16:39:19 ; Search time 78.4963 Seconds  
(without alignments)  
282.715 Million cell updates/sec

Title: US-10-775-481A-5  
Perfect score: 126  
Sequence: 1 NSSNYCCLCCNPACNGCY 19  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 6959266 seqs, 1168006243 residues  
Total number of hits satisfying chosen parameters: 6959266

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main.\*  
1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US097A\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US097B\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US099A\_COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US099B\_COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*  
28: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*  
29: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep.\*  
30: /cgn2\_6/ptodata/1/paa/US104\_COMB.pep.\*  
31: /cgn2\_6/ptodata/1/paa/US105\_COMB.pep.\*  
32: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep.\*  
33: /cgn2\_6/ptodata/1/paa/US107\_COMB.pep.\*  
34: /cgn2\_6/ptodata/1/paa/US108\_COMB.pep.\*  
35: /cgn2\_6/ptodata/1/paa/US109\_COMB.pep.\*  
36: /cgn2\_6/ptodata/1/paa/US110\_COMB.pep.\*  
37: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID                | Description       |
|------------|-------|-------------|--------|----|-------------------|-------------------|
| 1          | 126   | 100.0       | 19     | 1  | PCT-US04-03765-5  | Sequence 5, Appli |
| 2          | 126   | 100.0       | 19     | 1  | PCT-US94-12232-5  | Sequence 5, Appli |
| 3          | 126   | 100.0       | 19     | 8  | US-08-468-449B-5  | Sequence 5, Appli |
| 4          | 126   | 100.0       | 19     | 16 | US-09-263-477-5   | Sequence 5, Appli |
| 5          | 126   | 100.0       | 19     | 16 | US-09-263-477A-5  | Sequence 5, Appli |
| 6          | 126   | 100.0       | 19     | 21 | US-09-724-983-5   | Sequence 5, Appli |
| 7          | 126   | 100.0       | 19     | 32 | US-10-621-684-5   | Sequence 5, Appli |
| 8          | 126   | 100.0       | 19     | 33 | US-10-775-481A-5  | Sequence 5, Appli |
| 9          | 120   | 95.2        | 19     | 1  | PCT-US02-09551-23 | Sequence 23, Appl |
| 10         | 120   | 95.2        | 19     | 4  | US-08-085-126-20  | Sequence 20, Appl |
| 11         | 120   | 95.2        | 19     | 7  | US-08-342-241A-31 | Sequence 31, Appl |
| 12         | 120   | 95.2        | 19     | 8  | US-08-438-114-20  | Sequence 20, Appl |
| 13         | 120   | 95.2        | 19     | 16 | US-09-291-520-31  | Sequence 31, Appl |
| 14         | 120   | 95.2        | 19     | 16 | US-09-291-520-31  | Sequence 31, Appl |
| 15         | 120   | 95.2        | 19     | 19 | US-09-525-715-2   | Sequence 2, Appl  |
| 16         | 120   | 95.2        | 19     | 27 | US-10-107-814-23  | Sequence 23, Appl |
| 17         | 120   | 95.2        | 19     | 29 | US-10-371-966-1   | Sequence 1, Appli |
| 18         | 120   | 95.2        | 19     | 30 | US-10-479-606-7   | Sequence 7, Appli |
| 19         | 120   | 95.2        | 19     | 33 | US-10-766-735-1   | Sequence 1, Appli |
| 20         | 120   | 95.2        | 19     | 33 | US-10-766-735-26  | Sequence 26, Appl |
| 21         | 120   | 95.2        | 19     | 33 | US-10-796-719-1   | Sequence 1, Appli |
| 22         | 120   | 95.2        | 19     | 33 | US-10-796-719-26  | Sequence 26, Appl |
| 23         | 120   | 95.2        | 21     | 33 | US-10-766-735-39  | Sequence 39, Appl |
| 24         | 120   | 95.2        | 21     | 33 | US-10-796-719-39  | Sequence 39, Appl |
| 25         | 120   | 95.2        | 72     | 33 | US-10-766-735-21  | Sequence 21, Appl |
| 26         | 120   | 95.2        | 72     | 33 | US-10-796-719-21  | Sequence 21, Appl |
| 27         | 118   | 93.7        | 19     | 33 | US-10-766-735-84  | Sequence 84, Appl |
| 28         | 118   | 93.7        | 19     | 33 | US-10-766-735-86  | Sequence 86, Appl |
| 29         | 118   | 93.7        | 19     | 33 | US-10-796-719-84  | Sequence 84, Appl |
| 30         | 118   | 93.7        | 19     | 33 | US-10-796-719-86  | Sequence 86, Appl |
| 31         | 117   | 92.9        | 19     | 33 | US-10-766-735-92  | Sequence 92, Appl |
| 32         | 117   | 92.9        | 19     | 33 | US-10-796-719-92  | Sequence 92, Appl |
| 33         | 116   | 92.1        | 19     | 33 | US-10-371-966-2   | Sequence 2, Appli |
| 34         | 116   | 92.1        | 19     | 33 | US-10-766-735-27  | Sequence 27, Appl |
| 35         | 116   | 92.1        | 19     | 33 | US-10-766-735-87  | Sequence 87, Appl |
| 36         | 116   | 92.1        | 19     | 33 | US-10-796-719-27  | Sequence 27, Appl |
| 37         | 116   | 92.1        | 19     | 33 | US-10-796-719-87  | Sequence 87, Appl |
| 38         | 116   | 92.1        | 21     | 33 | US-10-766-735-40  | Sequence 40, Appl |
| 39         | 116   | 92.1        | 21     | 33 | US-10-766-735-41  | Sequence 41, Appl |
| 40         | 116   | 92.1        | 21     | 33 | US-10-796-719-40  | Sequence 40, Appl |
| 41         | 116   | 92.1        | 21     | 33 | US-10-796-719-41  | Sequence 41, Appl |
| 42         | 115   | 91.3        | 19     | 33 | US-10-766-735-28  | Sequence 28, Appl |
| 43         | 115   | 91.3        | 19     | 33 | US-10-766-735-75  | Sequence 75, Appl |
| 44         | 115   | 91.3        | 19     | 33 | US-10-766-735-79  | Sequence 79, Appl |
| 45         | 115   | 91.3        | 19     | 33 | US-10-766-735-90  | Sequence 90, Appl |

ALIGNMENTS

RESULT 1  
PCT-US04-03765-5  
; Sequence 5, Application PC/TUS0403765  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; APPLICANT: Pitari, Giovanni Mario  
; APPLICANT: Park, Jason  
; APPLICANT: Schulz, Stephanie  
; APPLICANT: Wolfe, Henry R.  
; APPLICANT: Lubbe, Wilhelm  
; TITLE OF INVENTION: The Use Of GCC Ligands  
; FILE REFERENCE: 08321-168 PCI  
; CURRENT APPLICATION NUMBER: PCT/US04/03765  
; CURRENT FILING DATE: 2004-02-10  
; PRIOR APPLICATION NUMBER: US 60/446,730  
; PRIOR FILING DATE: 2003-02-10  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5

; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: ST Ib  
PCT-US04-03765-5

Query Match 100.0%; Score 126; DB 1; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.7e-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NSSNYCCELCCNPACNGCY 19  
Db 1 NSSNYCCELCCNPACNGCY 19

## RESULT 2

PCT-US94-12232-5  
; Sequence 5, Application PC/TUS9412232  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: Compositions That Specifically  
; TITLE OF INVENTION: Bind To Colorectal Cancer Cells  
; TITLE OF INVENTION: And Methods Of Using The Same  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; ADDRESSEE: Norris  
; STREET: One Liberty Place, 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/12232  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/141,892  
; FILING DATE: 26-OCT-1993  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/305,056  
; FILING DATE: 13-SEP-1994  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: TJU-1360  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US94-12232-5

Query Match 100.0%; Score 126; DB 1; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.7e-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NSSNYCCELCCNPACNGCY 19  
Db 1 NSSNYCCELCCNPACNGCY 19

## RESULT 3

US-08-468-449B-5  
; Sequence 5, Application US/08468449B  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: ST Receptor Binding Compounds And Methods Of Using The Same  
; FILE REFERENCE: TJU-1588  
; CURRENT APPLICATION NUMBER: US/08/468,449B  
; CURRENT FILING DATE: 1995-06-06  
; PRIOR APPLICATION NUMBER: 08/141,892  
; PRIOR FILING DATE: 1993-10-26  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 5  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Novel Sequence  
US-08-468-449B-5

Query Match 100.0%; Score 126; DB 8; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.7e-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSSNYCCELCCNPACNGCY 19  
Db 1 NSSNYCCELCCNPACNGCY 19

## RESULT 4

US-09-263-477-5  
; Sequence 5, Application US/09263477  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: ST Receptor Binding Compounds and Methods  
; TITLE OF INVENTION: of Using the Same  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and Norris  
; STREET: One Liberty Place - 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: U.S.A.  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch disk, 720 Kb  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/263,477  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/141,892  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: TJU-0903  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-263-477-5

```
Query Match      100.0%; Score 126; DB 16; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSSNYCCCLCCNPACNGCY 19
Db 1 NSSNYCCCLCCNPACNGCY 19

RESULT 5
US-09-263-477A-5
; Sequence 5, Application US/09263477A
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; TITLE OF INVENTION: ST Receptor Binding Compounds and
; Methods of Using the Same
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: Wordperfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/263,477A
; FILING DATE: 05-Mar-1999
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,447A
; FILING DATE: 05-JAN-1996
; APPLICATION NUMBER: US 08/141,892
; FILING DATE: 26-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1702
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-263-477A-5

Query Match      100.0%; Score 126; DB 16; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSSNYCCCLCCNPACNGCY 19
Db 1 NSSNYCCCLCCNPACNGCY 19

RESULT 6
US-09-724-983-5
; Sequence 5, Application US/09724983
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; TITLE OF INVENTION: ST Receptor Binding Compounds and Methods of Using the Same
; FILE REFERENCE: TJU-2444
; CURRENT APPLICATION NUMBER: US/09/724,983
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 08/468,449

Query Match      100.0%; Score 126; DB 16; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSSNYCCCLCCNPACNGCY 19
Db 1 NSSNYCCCLCCNPACNGCY 19

RESULT 7
US-10-621-684-5
; Sequence 5, Application US/10621684
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; TITLE OF INVENTION: ST Receptor Binding Compounds and
; Methods of Using the Same
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: Wordperfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/621,684
; FILING DATE: 17-Jul-2003
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,447A
; FILING DATE: 05-JAN-1996
; APPLICATION NUMBER: US 08/141,892
; FILING DATE: 26-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1702
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-621-684-5

Query Match      100.0%; Score 126; DB 32; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSSNYCCCLCCNPACNGCY 19
Db 1 NSSNYCCCLCCNPACNGCY 19

RESULT 8
```

US-10-775-481A-5  
; Sequence 5, Application US/10775481A  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; APPLICANT: Pitari, Giovanni Mario  
; APPLICANT: Park, Jason  
; APPLICANT: Schulz, Stephanie  
; APPLICANT: Wolfe, Henry R.  
; APPLICANT: Lubbe, Wilhelm  
; TITLE OF INVENTION: The Use Of GCC Ligands  
; FILE REFERENCE: 08321-0168 US1  
; CURRENT APPLICATION NUMBER: US/10/775,481A  
; CURRENT FILING DATE: 2004-02-10  
; PRIOR APPLICATION NUMBER: US 60/446,730  
; PRIOR FILING DATE: 2003-02-10  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: heat stable toxin peptide Ib  
US-10-775-481A-5

Query Match 100.0%; Score 136; DB 33; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.7e-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPACNGCY 19  
|||||  
Db 1 NSSNYCCCLCCNPACNGCY 19

## RESULT 9

PCT-US02-09551-23  
; Sequence 23, Application PC/TUS0209551  
; GENERAL INFORMATION:  
; APPLICANT: SYNERGY PHARMACEUTICALS  
; TITLE OF INVENTION: GUANYLATE CYCLASE RECEPTOR AGONISTS FOR THE TREATMENT  
; FILE REFERENCE: 81361/141030  
; CURRENT APPLICATION NUMBER: PCT/US02/09551  
; CURRENT FILING DATE: 2002-03-28  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 23  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
; FEATURE:  
; NAME/KEY: DISULFID  
; LOCATION: (6)..(10)  
; FEATURE:  
; NAME/KEY: DISULFID  
; LOCATION: (7)..(15)  
; FEATURE:  
; NAME/KEY: DISULFID  
; LOCATION: (11)..(18)  
PCT-US02-09551-23

Query Match 95.2%; Score 120; DB 1; Length 19;  
Best Local Similarity 94.7%; Pred. No. 7.8e-06;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPACNGCY 19  
|||||  
Db 1 NSSNYCCCLCCNPACTGCGY 19

## RESULT 10

US-08-085-126-20  
; Sequence 20, Application US/08085126

; GENERAL INFORMATION:  
; APPLICANT: LADNER, Robert C.  
; APPLICANT: CANNON, Larick E.  
; TITLE OF INVENTION: CHELYDRYZYMES AND NEPAZYMES, AND USES  
; TITLE OF INVENTION: THEREFOR  
; NUMBER OF SEQUENCES: 236  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK  
; STREET: 419 Seventh Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/085,126  
; FILING DATE: 25-JUN-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: COOPER, Iver P.  
; REGISTRATION NUMBER: 28,005  
; REFERENCE/DOCKET NUMBER: LADNER=10  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-628-5197  
; TELEFAX: 202-737-3528  
; TELEX: 248633  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-085-126-20  
  
Query Match 95.2%; Score 120; DB 4; Length 19;  
Best Local Similarity 94.7%; Pred. No. 7.8e-06;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 NSSNYCCCLCCNPACNGCY 19  
|||||  
Db 1 NSSNYCCCLCCNPACTGCGY 19  
  
RESULT 11  
US-08-342-241A-31  
; Sequence 31, Application US/08342241A  
; GENERAL INFORMATION:  
; APPLICANT: Holmgren, Jan  
; APPLICANT: Lebens, Michael  
; TITLE OF INVENTION: IMMUNOGENS FOR STIMULATING MUCOSAL  
; TITLE OF INVENTION: IMMUNITY  
; NUMBER OF SEQUENCES: 51  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson and Bear  
; STREET: 620 Newport Center Drive 16th Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/342,241A  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:

APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: McCormack, Myra H  
REGISTRATION NUMBER: 36,602  
REFERENCE/DOCKET NUMBER: HOLMG.001A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176  
TELEX:  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acids  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
US-08-342-241A-31

Query Match 95.2%; Score 120; DB 7; Length 19;  
Best Local Similarity 94.7%; Pred. No. 7.8e-06;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NSSNYCCCLCCNPACNGCY 19  
Db 1 NSSNYCCCLCCNPACTGCT 19

## RESULT 12

US-08-438-114-20  
Sequence 20, Application US/08438114  
GENERAL INFORMATION:  
APPLICANT: LADNER, Richard C.  
APPLICANT: CANNON, Larick E.  
TITLE OF INVENTION: CHELYDRYZZYMS AND NEPAZYMS, AND USES  
TITLE OF INVENTION: THEREFOR  
NUMBER OF SEQUENCES: 240  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/438,114  
FILING DATE: 08-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/308,793  
FILING DATE: 19-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/085,126  
FILING DATE: 25-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: COOPER, Iver P.  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: LADNER=10B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:

LENGTH: 19 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-438-114-20

Query Match 95.2%; Score 120; DB 8; Length 19;  
Best Local Similarity 94.7%; Pred. No. 7.8e-06;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NSSNYCCCLCCNPACNGCY 19  
Db 1 NSSNYCCCLCCNPACTGCT 19

## RESULT 13

US-09-291-520-31  
Sequence 31, Application US/09291520  
GENERAL INFORMATION:  
APPLICANT: Holmgren, Jan  
APPLICANT: Lebens, Michael  
TITLE OF INVENTION: IMMUNOGENS FOR STIMULATING MUCOSAL  
TITLE OF INVENTION: IMMUNITY  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson and Bear  
STREET: 620 Newport Center Drive 16th Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/291,520  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/342,241  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: McCormack, Myra H  
REGISTRATION NUMBER: 36,602  
REFERENCE/DOCKET NUMBER: HOLMG.001A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176  
TELEX:  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acids  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
US-09-291-520-31

Query Match 95.2%; Score 120; DB 16; Length 19;  
Best Local Similarity 94.7%; Pred. No. 7.8e-06;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NSSNYCCCLCCNPACNGCY 19  
Db 1 NSSNYCCCLCCNPACTGCT 19

```

RESULT 14
US-09-291-520-31
; Sequence 31, Application US/09291520A
; GENERAL INFORMATION:
; APPLICANT: Holmgren, Jan
; APPLICANT: Lebens, Michael
; TITLE OF INVENTION: IMMUNOGENS FOR STIMULATING MUCOSAL
; TITLE OF INVENTION: IMMUNITY
; FILE REFERENCE: HOLMG.001DV1
; CURRENT APPLICATION NUMBER: US/09/291,520A
; CURRENT FILING DATE: 1999-04-04
; EARLIER APPLICATION NUMBER: 08/342,241
; EARLIER FILING DATE: 1994-11-17
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 19
; TYPE: PRT
; ORGANISM: E. coli
US-09-291-520-31

Query Match          95.2%; Score 120; DB 16; Length 19;
Best Local Similarity 94.7%; Pred. No. 7.8e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPACNGCY 19
Db 1 NSSNYCCCLCCNPACTGCG 19

RESULT 15
US-09-525-715-2
; Sequence 2, Application US/09525715
; GENERAL INFORMATION:
; APPLICANT: Johnson, David
; APPLICANT: Losco, Patricia
; APPLICANT: Harris, Mary Ann
; APPLICANT: Hagen, Else Kruger
; APPLICANT: Desai, Vinay
; TITLE OF INVENTION: Transplantable Sarcoma
; FILE REFERENCE: Transplantable Sarcoma-NIDN73110
; CURRENT APPLICATION NUMBER: US/09/525,715
; CURRENT FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/124,437
; PRIOR FILING DATE: 1999-03-15
; PRIOR APPLICATION NUMBER: GB9907408.0
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-525-715-2

Query Match          95.2%; Score 120; DB 19; Length 19;
Best Local Similarity 94.7%; Pred. No. 7.8e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPACNGCY 19
Db 1 NSSNYCCCLCCNPACTGCG 19

Search completed: March 26, 2005, 17:10:17
Job time : 79.4963 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2005, 16:44:35 ; Search time 5.89416 Seconds  
(without alignments)  
187.395 Million cell updates/sec

Title: US-10-775-481A-5

Perfect score: 126  
Sequence: 1 NSSNYCCELCCNPACNGCY 19

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 260697 seqs, 58133403 residues

Total number of hits satisfying chosen parameters: 260697

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents\_AA\_New.\*  
1: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US11\_NEW\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description          |
|------------|-------|-------------|--------|-------|----------------------|
| 1          | 98    | 77.8        | 18     | 6     | US-10-934-728-24     |
| 2          | 58    | 46.0        | 16     | 7     | US-11-066-697-1249   |
| 3          | 56    | 44.4        | 282    | 8     | US-60-655-875-128831 |
| 4          | 55    | 43.7        | 98     | 6     | US-10-489-448-1215   |
| 5          | 55    | 43.7        | 154    | 6     | US-10-489-448-1582   |
| 6          | 54.5  | 43.3        | 103    | 6     | US-10-489-448-1704   |
| 7          | 54    | 42.9        | 690    | 6     | US-10-938-061-106    |
| 8          | 54    | 42.9        | 690    | 6     | US-10-936-626-106    |
| 9          | 54    | 42.9        | 1071   | 1     | PCT-US05-05406-6     |
| 10         | 54    | 42.9        | 1071   | 7     | US-11-060-291-6      |
| 11         | 54    | 42.9        | 1123   | 6     | US-10-287-436A-285   |
| 12         | 54    | 42.9        | 3398   | 1     | PCT-US05-03560-6     |
| 13         | 52    | 41.3        | 152    | 8     | US-60-655-875-169174 |
| 14         | 52    | 41.3        | 158    | 6     | US-60-655-875-163386 |
| 15         | 52    | 41.3        | 1435   | 6     | US-10-184-644-581    |
| 16         | 52    | 41.3        | 1435   | 6     | US-10-192-007-581    |
| 17         | 51    | 40.5        | 111    | 6     | US-10-450-763-55164  |
| 18         | 50.5  | 40.1        | 271    | 8     | US-10-450-763-56787  |
| 19         | 50.5  | 40.1        | 271    | 8     | US-60-655-875-135230 |
| 20         | 50    | 39.7        | 15     | 6     | US-10-934-728-25     |
| 21         | 50    | 39.7        | 15     | 7     | US-11-066-697-1247   |
| 22         | 50    | 39.7        | 15     | 7     | US-11-066-697-1248   |
| 23         | 50    | 39.7        | 309    | 1     | PCT-US04-17965-1363  |
| 24         | 50    | 39.7        | 309    | 1     | PCT-US04-17965B-1363 |
| 25         | 50    | 39.7        | 358    | 8     | US-60-655-875-132098 |

|    |      |      |      |   |                      |                    |
|----|------|------|------|---|----------------------|--------------------|
| 26 | 50   | 39.7 | 424  | 6 | US-10-450-763-42525  | Sequence 42525, A  |
| 27 | 50   | 39.7 | 508  | 6 | US-10-489-448-1646   | Sequence 1646, Ap  |
| 28 | 49.5 | 39.3 | 206  | 6 | US-10-450-763-59727  | Sequence 59727, A  |
| 29 | 49.5 | 39.3 | 1200 | 7 | US-11-058-476-3      | Sequence 3, Appli  |
| 30 | 49   | 38.9 | 12   | 6 | US-10-934-728-22     | Sequence 22, Appli |
| 31 | 49   | 38.9 | 162  | 8 | US-60-655-875-165998 | Sequence 165998,   |
| 32 | 49   | 38.9 | 264  | 8 | US-60-655-875-137235 | Sequence 137235,   |
| 33 | 49   | 38.9 | 282  | 8 | US-60-655-875-139718 | Sequence 139718,   |
| 34 | 49   | 38.9 | 857  | 8 | US-60-655-875-153943 | Sequence 153943,   |
| 35 | 48.5 | 38.5 | 256  | 6 | US-10-450-763-38696  | Sequence 38696, A  |
| 36 | 48.5 | 38.5 | 351  | 6 | US-10-450-763-54057  | Sequence 54057, A  |
| 37 | 48.5 | 38.5 | 398  | 6 | US-10-450-763-49367  | Sequence 49367, A  |
| 38 | 48.5 | 38.5 | 497  | 6 | US-10-287-436A-269   | Sequence 269, App  |
| 39 | 48.5 | 38.5 | 1501 | 5 | US-09-857-581B-16    | Sequence 16, Appl  |
| 40 | 48   | 38.1 | 13   | 7 | US-11-066-697-1392   | Sequence 1392, Ap  |
| 41 | 48   | 38.1 | 156  | 8 | US-60-655-875-162472 | Sequence 162472,   |
| 42 | 48   | 38.1 | 181  | 8 | US-60-655-875-169804 | Sequence 169804,   |
| 43 | 48   | 38.1 | 212  | 8 | US-60-655-875-169914 | Sequence 169914,   |
| 44 | 48   | 38.1 | 221  | 8 | US-60-655-875-162864 | Sequence 162864,   |
| 45 | 48   | 38.1 | 560  | 6 | US-10-287-436A-220   | Sequence 220, App  |

ALIGNMENTS

RESULT 1  
US-10-934-728-24  
; Sequence 24, Application US/10934728  
; GENERAL INFORMATION:  
; APPLICANT: Balloul, Jean Marc  
; APPLICANT: Paul, Stephane  
; APPLICANT: Geist, Michel  
; APPLICANT: Silvestre, Nathalie  
; APPLICANT: Erba, Philippe  
; TITLE OF INVENTION: Foxvirus With Targeted Infection Specificity  
; FILE REFERENCE: 032751-115  
; CURRENT APPLICATION NUMBER: US/10/934,728  
; PRIOR FILING DATE: 2004-09-07  
; PRIOR APPLICATION NUMBER: US 09/832,899  
; PRIOR FILING DATE: 2001-04-14  
; PRIOR APPLICATION NUMBER: US 60/246,080  
; PRIOR FILING DATE: 2000-11-07  
; PRIOR APPLICATION NUMBER: EP 0040109.7  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: EP 01440009.7  
; PRIOR FILING DATE: 2001-01-22  
; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 24  
; LENGTH: 18  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Sta ligand  
US-10-934-728-24

Query Match 77.8% Score 98; DB 6; Length 18;  
Best Local Similarity 93.3% Pred.No. 1.7e-06;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 5 YCCELCCNPACNGCY 19  
| | | | | | | | | | | | | | | | | |  
Db 4 YCCELCCNPACAGCY 18

RESULT 2  
US-11-066-697-1249  
; Sequence 1249, Application US/110666697  
; GENERAL INFORMATION:  
; APPLICANT: Bridon, Dominique P.  
; APPLICANT: Exrin, Alan M.  
; APPLICANT: Milner, Peter G.  
; APPLICANT: Holmes, Darren L.

APPLICANT: Thibaudau, Karen  
; TITLE OF INVENTION: PROTECTION OF ENDOGENOUS THERAPEUTIC PEPTIDES FROM  
; TITLE OF INVENTION: PEPTIDASE ACTIVITY THROUGH CONJUGATION TO BLOOD  
; TITLE OF INVENTION: COMPONENTS  
; FILE REFERENCE: 500862002301  
; CURRENT APPLICATION NUMBER: US/11/066,697  
; CURRENT FILING DATE: 2005-02-25  
; PRIOR APPLICATION NUMBER: 09/657,276  
; PRIOR FILING DATE: 2000-09-07  
; PRIOR APPLICATION NUMBER: 60/153,406  
; PRIOR FILING DATE: 1999-09-10  
; PRIOR APPLICATION NUMBER: 60/159,783  
; PRIOR FILING DATE: 1999-10-15  
; NUMBER OF SEQ ID NOS: 1617  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1249  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Peptide  
US-11-066-697-1249

Query Match 46.0%; Score 58; DB 7; Length 16;  
Best Local Similarity 75.0%; Pred. No. 0.15;  
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 CELCCNPACNGC 18  
|||||  
Db 4 CELCVNACTGC 15  
|||||

RESULT 3  
US-60-655-875-128831  
; Sequence 128831, Application US/60655875  
; GENERAL INFORMATION:  
; APPLICANT: Boukharov, Andrey  
; APPLICANT: Du, Zijing  
; APPLICANT: Guo, Liang  
; APPLICANT: Kovalic, David  
; APPLICANT: Lu, Maolong  
; APPLICANT: McCarter, James  
; APPLICANT: Miller, Nancy  
; APPLICANT: Williams, Deryck  
; APPLICANT: Vaudin, Mark  
; APPLICANT: Wu, Wei  
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS  
; TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF  
; FILE REFERENCE: 38-21(53885)  
; CURRENT APPLICATION NUMBER: US/60/655,875  
; CURRENT FILING DATE: 2005-02-24  
; NUMBER OF SEQ ID NOS: 171306  
; SEQ ID NO 128831  
; LENGTH: 282  
; TYPE: PRT  
; ORGANISM: Heterodera glycines  
; FEATURE:  
; OTHER INFORMATION: Coding regions on vcdna: vcdna-SeqID\_55254; Strand=-; Position=1-  
; FEATURE:  
; FEATURE:  
; OTHER INFORMATION: Homolog annotation: Hit ID=XP\_311331.1; Match level="QueryCoverag  
; OTHER INFORMATION: =100%; HitCoverage=100%; E-value=5e-34, Identity=36%; Hit  
; OTHER INFORMATION: description=ENSANGP00000001657 [Anopheles gambiae]  
US-60-655-875-128831

Query Match 44.4%; Score 56; DB 8; Length 282;  
Best Local Similarity 52.6%; Pred. No. 3.3;  
Matches 10; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

QY 2 SSNVCC--ELCCNPACNGC 18  
|||||  
Db 239 SSNHCCTSSSCCNHSSNRC 257  
|||||

RESULT 4  
US-10-489-448-3215  
; Sequence 3215, Application US/10489448  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ghosh, Malabika  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Ma, Yunqing  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: Weng, Gezhi  
; APPLICANT: Haley-Vicente, Dana  
; APPLICANT: Dmanac, Radoje T  
; TITLE OF INVENTION: Novel Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 810CIP PCT  
; CURRENT APPLICATION NUMBER: US/10/489,448  
; CURRENT FILING DATE: 1004-03-10  
; PRIOR APPLICATION NUMBER: US 60/324,631  
; PRIOR FILING DATE: 2001-09-24  
; PRIOR APPLICATION NUMBER: US 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; PRIOR APPLICATION NUMBER: US 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: PCT/US00/35017  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/491,404  
; PRIOR FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: PCT/US01/02623  
; PRIOR FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: US 09/496,914  
; PRIOR FILING DATE: 2000-02-03  
; PRIOR APPLICATION NUMBER: US 09/560,875  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: PCT/US01/03800  
; PRIOR FILING DATE: 2001-02-05  
; PRIOR APPLICATION NUMBER: US 09/515,126  
; PRIOR FILING DATE: 2000-02-28  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 3476  
; SOFTWARE: pt\_FL\_genes Version 6.0  
; SEQ ID NO 3215  
; LENGTH: 98  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(98)  
; OTHER INFORMATION: Xaa = any amino acid or symbol as shown in the table 8 as set fort  
US-10-489-448-3215

Query Match 43.7%; Score 55; DB 6; Length 98;  
Best Local Similarity 70.0%; Pred. No. 1.7;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 CCELCNCPAC 15  
|||||  
Db 16 CCQPCCRPAC 25  
|||||

RESULT 5  
US-10-489-448-1582  
; Sequence 1582, Application US/10489448  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom



```

; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wang, Dunrui
; APPLICANT: Ma, Yunqing
; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Zhiwei
; APPLICANT: Haley-Vicente, Dana
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/10/489,448
; CURRENT FILING DATE: 1004-03-10
; PRIOR APPLICATION NUMBER: US 60/324,631
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: PCT/US00/35017
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: PCT/US01/03800
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 3476
; SOFTWARE: pc_FL_genes Version 6.0
; SEQ ID NO 1582
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-489-448-1582

```

```

Query Match 43.7%; Score 55; DB 6; Length 154;
Best Local Similarity 70.0%; Pred. No. 2.6;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 6 CCELCNCPAC 15
Db 125 CCQPCCRPAC 134

```

```

RESULT 6
US-10-489-448-1704
; Sequence 1704, Application US/10489448
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wang, Dunrui
; APPLICANT: Ma, Yunqing

```

```

; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Zhiwei
; APPLICANT: Weng, Gezhi
; APPLICANT: Haley-Vicente, Dana
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/10/489,448
; CURRENT FILING DATE: 1004-03-10
; PRIOR APPLICATION NUMBER: US 60/324,631
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: PCT/US00/35017
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: PCT/US01/03800
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 3476
; SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 1704
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-489-448-1704

```

```

Query Match 43.3%; Score 54.5; DB 6; Length 103;
Best Local Similarity 61.5%; Pred. No. 2.1;
Matches 8; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

```

```

Qy 6 CCELCNCPACNGC 18
Db 45 CCSSCC-PCCRCG 56

```

```

RESULT 7
US-10-938-061-106
; Sequence 106, Application US/10938061
; GENERAL INFORMATION:
; APPLICANT: Cairns, Belinda
; APPLICANT: Chen, Ruihuan
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Koeppe, Hartmut
; APPLICANT: Phillips, Heidi S.
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan D.
; APPLICANT: Smith, Victoria
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas D.
; APPLICANT: Zhang, Zemin
; APPLICANT: Sakanaka, Chie
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Reed Chae J.
; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; FILE REFERENCE: P5001R1P1B
; CURRENT APPLICATION NUMBER: US/10/938,061
; CURRENT FILING DATE: 2004-09-10
; PRIOR APPLICATION NUMBER: US 10/872,991

```

; PRIOR FILING DATE: 2004-06-21  
; PRIOR APPLICATION NUMBER: US 10/872,972  
; PRIOR FILING DATE: 2004-06-21  
; PRIOR APPLICATION NUMBER: US 10/241,220  
; PRIOR FILING DATE: 2002-09-11  
; PRIOR APPLICATION NUMBER: US 10/177,488  
; PRIOR FILING DATE: 2002-06-19  
; PRIOR APPLICATION NUMBER: US 60/299,500  
; PRIOR FILING DATE: 2001-06-20  
; PRIOR APPLICATION NUMBER: US 60/301,880  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/323,268  
; PRIOR FILING DATE: 2001-09-18  
; PRIOR APPLICATION NUMBER: US 60/557,116  
; PRIOR FILING DATE: 2004-03-26  
; PRIOR APPLICATION NUMBER: US 60/598,899  
; PRIOR FILING DATE: 2004-08-04  
; NUMBER OF SEQ ID NOS: 154  
; SEQ ID NO 106  
; LENGTH: 690  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-938-061-106

Query Match 42.9%; Score 54; DB 6; Length 690;  
Best Local Similarity 50.0%; Pred. No. 13;  
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 5 YCCELCNPNACNGC 18  
|||:|:|  
Db 620 YCCRVCCACCLLC 633

RESULT 8  
US-10-936-626-106  
; Sequence 106, Application US/10936626  
; GENERAL INFORMATION:  
; APPLICANT: Cairns, Belinda  
; APPLICANT: Chen, Ruihuan  
; APPLICANT: Frantz, Gretchen  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Koepfen, Hartmut  
; APPLICANT: Phillips, Heidi S.  
; APPLICANT: Polakis, Paul  
; APPLICANT: Spencer, Susan D.  
; APPLICANT: Smith, Victoria  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wu, Thomas D.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and  
; FILE REFERENCE: P5001R1P1  
; CURRENT APPLICATION NUMBER: US/10/936,626  
; CURRENT FILING DATE: 2004-09-08  
; PRIOR APPLICATION NUMBER: US 10/872,991  
; PRIOR FILING DATE: 2004-06-21  
; PRIOR APPLICATION NUMBER: US 10/872,972  
; PRIOR FILING DATE: 2004-06-21  
; PRIOR APPLICATION NUMBER: US 10/241,220  
; PRIOR FILING DATE: 2002-09-11  
; PRIOR APPLICATION NUMBER: US 10/177,488  
; PRIOR FILING DATE: 2002-06-19  
; PRIOR APPLICATION NUMBER: US 60/299,500  
; PRIOR FILING DATE: 2001-06-20  
; PRIOR APPLICATION NUMBER: US 60/301,880  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/323,268  
; PRIOR FILING DATE: 2001-09-18  
; PRIOR APPLICATION NUMBER: US 60/557,116  
; PRIOR FILING DATE: 2004-03-26  
; PRIOR APPLICATION NUMBER: US 60/598,899  
; PRIOR FILING DATE: 2004-08-04  
; NUMBER OF SEQ ID NOS: 154

; SEQ ID NO 106  
; LENGTH: 690  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-936-626-106

Query Match 42.9%; Score 54; DB 6; Length 690;  
Best Local Similarity 50.0%; Pred. No. 13;  
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 5 YCCELCNPNACNGC 18  
|||:|:|  
Db 620 YCCRVCCACCLLC 633

RESULT 9  
PCT-US05-05406-6  
; Sequence 6, Application PC/TUS0505406  
; GENERAL INFORMATION:  
; APPLICANT: Olsen, Byron  
; TITLE OF INVENTION: A Method for the Production of Transgenic Proteins Useful in the  
; FILE REFERENCE: GTC-223 PCT  
; CURRENT APPLICATION NUMBER: PCT/US05/05406  
; CURRENT FILING DATE: 2005-02-18  
; PRIOR APPLICATION NUMBER: 60/545,790  
; PRIOR FILING DATE: 2004-02-19  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 6  
; LENGTH: 1071  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US05-05406-6

Query Match 42.9%; Score 54; DB 1; Length 1071;  
Best Local Similarity 53.8%; Pred. No. 19;  
Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 6 CCELCNPNACNGC 18  
|||:|:|  
Db 261 CCGACCTCTCGGC 273

RESULT 10  
US-11-060-291-6  
; Sequence 6, Application US/11060291  
; GENERAL INFORMATION:  
; APPLICANT: Olsen, Byron  
; TITLE OF INVENTION: A Method for the Production of Transgenic Proteins Useful in the  
; FILE REFERENCE: GTC-223  
; CURRENT APPLICATION NUMBER: US/11/060,291  
; CURRENT FILING DATE: 2005-02-17  
; PRIOR APPLICATION NUMBER: 60/545,790  
; PRIOR FILING DATE: 2004-02-19  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 6  
; LENGTH: 1071  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-060-291-6

Query Match 42.9%; Score 54; DB 7; Length 1071;  
Best Local Similarity 53.8%; Pred. No. 19;  
Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 6 CCELCNPNACNGC 18  
|||:|:|  
Db 261 CCGACCTCTCGGC 273

RESULT 11  
US-10-287-436A-285  
; Sequence 285, Application US/10287436A  
; GENERAL INFORMATION:  
; APPLICANT: CHILDREN'S HOSPITAL MEDICAL CENTER  
; TITLE OF INVENTION: METHOD FOR DIAGNOSIS AND TREATMENT OF  
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS  
; FILE REFERENCE: 10872-514696  
; CURRENT APPLICATION NUMBER: US/10/287,436A  
; CURRENT FILING DATE: 2002-10-31  
; PRIOR APPLICATION NUMBER: US 60/336,220  
; PRIOR FILING DATE: 2001-10-31  
; NUMBER OF SEQ ID NOS: 1446  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 285  
; LENGTH: 1123  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-287-436A-285

Query Match 42.9%; Score 54; DB 6; Length 1123;  
Best Local Similarity 53.8%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
  
Qy 6 CCELCNPNACNGC 18  
Db 64 CCAGCCTTGCTGC 76

RESULT 12  
PCT-US05-03560-6  
; Sequence 6, Application PC/TUS0503560  
; GENERAL INFORMATION:  
; APPLICANT: EXELIXIS, INC.  
; TITLE OF INVENTION: ITPKS AS SMODIFIERS OF THE IGFR PATHWAY AND METHODS OF USE  
; FILE REFERENCE: EX05-004C-PC  
; CURRENT APPLICATION NUMBER: PCT/US05/03560  
; CURRENT FILING DATE: 2005-01-27  
; PRIOR APPLICATION NUMBER: US60/539,837  
; PRIOR FILING DATE: 2004-01-28  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 6  
; LENGTH: 3398  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US05-03560-6

Query Match 42.9%; Score 54; DB 1; Length 3398;  
Best Local Similarity 53.8%; Pred. No. 54;  
Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
  
Qy 6 CCELCNPNACNGC 18  
Db 2463 CCTCCCTGTCAGC 2475

RESULT 13  
US-60-655-875-169174  
; Sequence 169174, Application US/60655875  
; GENERAL INFORMATION:  
; APPLICANT: Boukharov, Andrey  
; APPLICANT: Du, Zijing  
; APPLICANT: Guo, Liang  
; APPLICANT: Kovalic, David  
; APPLICANT: Lu, Maolong  
; APPLICANT: McCarter, James  
; APPLICANT: Miller, Nancy  
; APPLICANT: Williams, Deryck  
; APPLICANT: Vaudin, Mark  
; APPLICANT: Wu, Wei  
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS  
; FILE REFERENCE: 38-21(53885)  
; CURRENT APPLICATION NUMBER: US/60/655,875  
; CURRENT FILING DATE: 2005-02-24  
; NUMBER OF SEQ ID NOS: 171306  
; SEQ ID NO 169174  
; LENGTH: 152  
; TYPE: PRT  
; ORGANISM: Heterodera glycines  
; FEATURE:  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
; NAME/KEY: misc feature  
; LOCATION: (4)..(4)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (14)..(14)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (18)..(21)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
; FEATURE:  
; OTHER INFORMATION: Coding regions on vcdna: vcdna=seqID\_89809; Strands=-; Position=0-  
US-60-655-875-169174

Query Match 41.3%; Score 52; DB 8; Length 158;  
Best Local Similarity 53.8%; Pred. No. 6.3;  
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
  
Qy 4 NYCCELCCNPACN 16  
Db 41 NICCRCCNCCCS 53

; FILE REFERENCE: 38-21(53885)  
; CURRENT APPLICATION NUMBER: US/60/655,875  
; CURRENT FILING DATE: 2005-02-24  
; NUMBER OF SEQ ID NOS: 171306  
; SEQ ID NO 169174  
; LENGTH: 152  
; TYPE: PRT  
; ORGANISM: Heterodera glycines  
; FEATURE:  
; OTHER INFORMATION: Coding regions on vcdna: vcdna=seqID\_95597; Strands=-; Position=3-  
US-60-655-875-169174

Query Match 41.3%; Score 52; DB 8; Length 152;  
Best Local Similarity 53.8%; Pred. No. 6.1;  
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
  
Qy 4 NYCCELCCNPACN 16  
Db 11 NICCRCCNCCCS 23

RESULT 14  
US-60-655-875-163386  
; Sequence 163386, Application US/60655875  
; GENERAL INFORMATION:  
; APPLICANT: Boukharov, Andrey  
; APPLICANT: Du, Zijing  
; APPLICANT: Guo, Liang  
; APPLICANT: Kovalic, David  
; APPLICANT: Lu, Maolong  
; APPLICANT: McCarter, James  
; APPLICANT: Miller, Nancy  
; APPLICANT: Williams, Deryck  
; APPLICANT: Vaudin, Mark  
; APPLICANT: Wu, Wei  
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS  
; FILE REFERENCE: 38-21(53885)  
; CURRENT APPLICATION NUMBER: US/60/655,875  
; CURRENT FILING DATE: 2005-02-24  
; NUMBER OF SEQ ID NOS: 171306  
; SEQ ID NO 163386  
; LENGTH: 158  
; TYPE: PRT  
; ORGANISM: Heterodera glycines  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (2)..(2)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
; NAME/KEY: misc feature  
; LOCATION: (4)..(4)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (14)..(14)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (18)..(21)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
; FEATURE:  
; OTHER INFORMATION: Coding regions on vcdna: vcdna=seqID\_89809; Strands=-; Position=0-  
US-60-655-875-163386

Query Match 41.3%; Score 52; DB 8; Length 158;  
Best Local Similarity 53.8%; Pred. No. 6.3;  
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
  
Qy 4 NYCCELCCNPACN 16  
Db 41 NICCRCCNCCCS 53

Mon Mar 28 09:44:30 2005

```

RESULT 15
US-10-184-644-581
; Sequence 581, Application US/10184644
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC227
; CURRENT APPLICATION NUMBER: US/10/184,644
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 581
; LENGTH: 1435
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-644-581

```

```

Query Match 41.3%; Score 52; DB 6; Length 1435;
Best Local Similarity 53.8%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 6 CCELCNPNACNGC 18
Db 1228 CCAACCACTAC 1240

```

Search completed: March 26, 2005, 17:11:53  
Job time : 6.89416 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model  
Run on: March 26, 2005, 16:39:19 ; Search time 437.927 Seconds  
(without alignments)  
282.715 Million cell updates/sec

Title: US-10-775-481A-55  
Perfect score: 548  
Sequence: 1 MSGQLMAVLLVLSAQ.....LRTIATDCELCINVACTGC 106

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 6959266 seqs, 116806243 residues

Total number of hits satisfying chosen parameters: 6959266

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents AA Main:  
1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US080\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US097A\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US097B\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US099A\_COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US099B\_COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*  
28: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*  
29: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep.\*  
30: /cgn2\_6/ptodata/1/paa/US104\_COMB.pep.\*  
31: /cgn2\_6/ptodata/1/paa/US105\_COMB.pep.\*  
32: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep.\*  
33: /cgn2\_6/ptodata/1/paa/US107\_COMB.pep.\*  
34: /cgn2\_6/ptodata/1/paa/US108\_COMB.pep.\*  
35: /cgn2\_6/ptodata/1/paa/US109\_COMB.pep.\*  
36: /cgn2\_6/ptodata/1/paa/US110\_COMB.pep.\*  
37: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID                   | Description       |
|------------|-------|-------|--------|----|----------------------|-------------------|
| 1          | 548   | 100.0 | 106    | 1  | PCT-US04-03765-55    | Sequence 55, Appl |
| 2          | 548   | 100.0 | 106    | 33 | US-10-775-481A-55    | Sequence 55, Appl |
| 3          | 358   | 65.3  | 111    | 37 | US-60-248-592-71     | Sequence 71, Appl |
| 4          | 358   | 65.3  | 122    | 37 | US-60-245-228-223    | Sequence 223, App |
| 5          | 356   | 65.0  | 112    | 1  | PCT-US04-03765-56    | Sequence 56, Appl |
| 6          | 356   | 65.0  | 112    | 1  | PCT-US04-36404-145   | Sequence 145, App |
| 7          | 356   | 65.0  | 112    | 27 | US-10-170-205B-19124 | Sequence 19124, A |
| 8          | 356   | 65.0  | 112    | 30 | US-10-479-606-5      | Sequence 5, Appli |
| 9          | 356   | 65.0  | 112    | 33 | US-10-700-439-145    | Sequence 145, App |
| 10         | 356   | 65.0  | 112    | 33 | US-10-775-481A-56    | Sequence 56, Appl |
| 11         | 356   | 65.0  | 112    | 37 | US-60-453-050-12227  | Sequence 12227, A |
| 12         | 356   | 65.0  | 112    | 37 | US-60-453-135-12227  | Sequence 12227, A |
| 13         | 356   | 65.0  | 112    | 37 | US-60-466-412-12227  | Sequence 12227, A |
| 14         | 356   | 65.0  | 112    | 37 | US-60-568-073-1020   | Sequence 1020, Ap |
| 15         | 306   | 55.8  | 93     | 23 | US-09-834-366-26398  | Sequence 26398, A |
| 16         | 306   | 55.8  | 93     | 37 | US-60-197-873-26398  | Sequence 26398, A |
| 17         | 252.5 | 46.1  | 109    | 30 | US-10-479-606-6      | Sequence 6, Appli |
| 18         | 166.5 | 30.4  | 115    | 1  | PCT-US04-18751-70    | Sequence 70, Appl |
| 19         | 166.5 | 30.4  | 115    | 23 | US-09-834-366-17911  | Sequence 17911, A |
| 20         | 166.5 | 30.4  | 115    | 25 | US-09-881-353-61     | Sequence 61, Appl |
| 21         | 166.5 | 30.4  | 115    | 26 | US-10-003-608-22     | Sequence 22, Appl |
| 22         | 166.5 | 30.4  | 115    | 27 | US-10-170-205B-19125 | Sequence 19125, A |
| 23         | 166.5 | 30.4  | 115    | 28 | US-10-235-994-22     | Sequence 22, Appl |
| 24         | 166.5 | 30.4  | 115    | 28 | US-10-262-473-12     | Sequence 12, Appl |
| 25         | 166.5 | 30.4  | 115    | 30 | US-10-479-606-4      | Sequence 4, Appli |
| 26         | 166.5 | 30.4  | 115    | 37 | US-60-197-873-17911  | Sequence 17911, A |
| 27         | 166.5 | 30.4  | 115    | 37 | US-60-239-841-61     | Sequence 61, Appl |
| 28         | 166.5 | 30.4  | 115    | 37 | US-60-453-050-12228  | Sequence 12228, A |
| 29         | 166.5 | 30.4  | 115    | 37 | US-60-453-135-12228  | Sequence 12228, A |
| 30         | 166.5 | 30.4  | 115    | 37 | US-60-466-412-12228  | Sequence 12228, A |
| 31         | 158   | 28.8  | 89     | 37 | US-60-177-670-190    | Sequence 190, App |
| 32         | 122.5 | 27.8  | 94     | 1  | PCT-US04-18751-68    | Sequence 68, Appl |
| 33         | 120   | 21.9  | 102    | 1  | PCT-US04-18751-71    | Sequence 71, Appl |
| 34         | 120   | 21.9  | 102    | 28 | US-10-262-473-14     | Sequence 14, Appl |
| 35         | 120   | 21.9  | 108    | 1  | PCT-US04-18751-72    | Sequence 72, Appl |
| 36         | 120   | 21.9  | 108    | 28 | US-10-262-473-16     | Sequence 16, Appl |
| 37         | 114   | 20.8  | 24     | 1  | PCT-US04-18751-77    | Sequence 77, Appl |
| 38         | 99    | 18.1  | 119    | 22 | US-09-791-537-68439  | Sequence 68439, A |
| 39         | 98.5  | 18.0  | 23     | 1  | PCT-US04-18751-78    | Sequence 78, Appl |
| 40         | 97.5  | 17.8  | 87     | 1  | PCT-US00-05883-1162  | Sequence 1162, Ap |
| 41         | 97.5  | 17.8  | 87     | 24 | US-09-925-299-1162   | Sequence 1162, Ap |
| 42         | 90    | 16.4  | 15     | 1  | PCT-US04-18751-48    | Sequence 48, Appl |
| 43         | 89    | 16.2  | 17     | 37 | US-60-160-203-3643   | Sequence 3643, Ap |
| 44         | 84    | 15.3  | 16     | 1  | PCT-US02-09551-20    | Sequence 20, Appl |
| 45         | 84    | 15.3  | 16     | 1  | PCT-US04-18751-73    | Sequence 73, Appl |

ALIGNMENTS

RESULT 1  
PCT-US04-03765-55  
; Sequence 55, Application PC/TUS0403765  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; APPLICANT: Pitari, Giovanni Mario  
; APPLICANT: Park, Jason  
; APPLICANT: Schulz, Stephanie  
; APPLICANT: Wolfe, Henry R.  
; APPLICANT: Lubbe, Wilhelm  
; TITLE OF INVENTION: The Use Of GCC Ligands  
; FILE REFERENCE: 08321-168 PCI  
; CURRENT APPLICATION NUMBER: PCT/US04/03765  
; CURRENT FILING DATE: 2004-02-10  
; PRIOR APPLICATION NUMBER: US 60/446,730  
; PRIOR FILING DATE: 2003-02-10  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 55

```
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Rattus norvegicus
PCT-US04-03765-55

Query Match      100.0%; Score 548; DB 1; Length 106;
Best Local Similarity 100.0%; Pred. No. 2.9e-54;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGSQLWAALLLVLSAQGVYIKYHGFQVQLESVKKLNLEEEKOMSDPQQQKSGLLPD 60
Db 1 MSGSQLWAALLLVLSAQGVYIKYHGFQVQLESVKKLNLEEEKOMSDPQQQKSGLLPD 60

QY 61 VCYNPALPLDLPVCASQEAASFTFKALRTIATDECELCINVACTGC 106
Db 61 VCYNPALPLDLPVCASQEAASFTFKALRTIATDECELCINVACTGC 106

RESULT 2
US-10-775-481A-55
; Sequence 55, Application US/10775481A
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; APPLICANT: Pitari, Giovanni Mario
; APPLICANT: Park, Jason
; APPLICANT: Schulz, Stephanie
; APPLICANT: Wolfe, Henry R.
; APPLICANT: Lubbe, Wilhelm
; TITLE OF INVENTION: The Use Of GCC Ligands
; FILE REFERENCE: 08321-0168 US1
; CURRENT APPLICATION NUMBER: US/10/775, 481A
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/446,730
; PRIOR FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-775-481A-55

Query Match      100.0%; Score 548; DB 33; Length 106;
Best Local Similarity 100.0%; Pred. No. 2.9e-54;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGSQLWAALLLVLSAQGVYIKYHGFQVQLESVKKLNLEEEKOMSDPQQQKSGLLPD 60
Db 1 MSGSQLWAALLLVLSAQGVYIKYHGFQVQLESVKKLNLEEEKOMSDPQQQKSGLLPD 60

QY 61 VCYNPALPLDLPVCASQEAASFTFKALRTIATDECELCINVACTGC 106
Db 61 VCYNPALPLDLPVCASQEAASFTFKALRTIATDECELCINVACTGC 106

RESULT 3
US-60-248-592-71
; Sequence 71, Application US/60248592
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN CYCLASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN CYCLASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000961
; CURRENT APPLICATION NUMBER: US/60/248,592
; CURRENT FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 71
; LENGTH: 111
; TYPE: PRT
; ORGANISM: HUMAN
US-60-248-592-71

Query Match      65.3%; Score 358; DB 37; Length 111;
Best Local Similarity 65.7%; Pred. No. 2.2e-32;
Matches 69; Conservative 14; Mismatches 22; Indels 0; Gaps 0;

QY 2 SGSQLWAALLLVLSAQGVYIKYHGFQVQLESVKKLNLEEEKOMSDPQQQKSGLLPDV 61
Db 7 SGLLTGVAALLLVLSAQGVYIKYHGFQVQLESVKKLNLEEEKOMSDPQQQKSGLLPDV 66

QY 62 CYNPALPLDLPVCASQEAASFTFKALRTIATDECELCINVACTGC 106
Db 67 CHHPALPDLPVCASQEAASFTFKALRTIATDECELCINVACTGC 111

RESULT 4
US-60-245-228-223
; Sequence 223, Application US/60245228
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN CYCLASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN CYCLASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000878
; CURRENT APPLICATION NUMBER: US/60/245,228
; CURRENT FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 630
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 223
; LENGTH: 122
; TYPE: PRT
; ORGANISM: HUMAN
US-60-245-228-223

Query Match      65.3%; Score 358; DB 37; Length 122;
Best Local Similarity 65.7%; Pred. No. 2.5e-32;
Matches 69; Conservative 14; Mismatches 22; Indels 0; Gaps 0;

QY 2 SGSQLWAALLLVLSAQGVYIKYHGFQVQLESVKKLNLEEEKOMSDPQQQKSGLLPDV 61
Db 17 SGLLTGVAALLLVLSAQGVYIKYHGFQVQLESVKKLNLEEEKOMSDPQQQKSGLLPDV 76

QY 62 CYNPALPLDLPVCASQEAASFTFKALRTIATDECELCINVACTGC 106
Db 77 CHHPALPDLPVCASQEAASFTFKALRTIATDECELCINVACTGC 121

RESULT 5
PCT-US04-03765-56
; Sequence 56, Application PC/TUS0403765
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; APPLICANT: Pitari, Giovanni Mario
; APPLICANT: Park, Jason
; APPLICANT: Schulz, Stephanie
; APPLICANT: Wolfe, Henry R.
; APPLICANT: Lubbe, Wilhelm
; TITLE OF INVENTION: The Use Of GCC Ligands
; FILE REFERENCE: 08321-168 PC1
; CURRENT APPLICATION NUMBER: PCT/US04/03765
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/446,730
; PRIOR FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US04-03765-56

Query Match      65.0%; Score 356; DB 1; Length 112;
Best Local Similarity 65.7%; Pred. No. 3.8e-32;
Matches 69; Conservative 14; Mismatches 22; Indels 0; Gaps 0;
```

|    |    |   |     |
|----|----|---|-----|
| QY | 2  | SGSOLMAVLLILLVLQSAQCQVYIKYHGFQVLESVKKLNLEEKQMSDPQQOKSGILPVP | 61  |
| D8 | 7  | SGLLPGVAVLLLLQLSTQSVYIQYGRFVLESMMKLSDLEAQWAPSRLQAQSILPAV    | 66  |
| QY | 62 | CYNPALDLDQPVCASQEAASTPKALRTATNDCELCINVACTGC                 | 106 |
| D8 | 67 | CHHPALPDLDQPVCASQEASSIEFKTURTANDDCELQNVACTGC                | 111 |

```

RESULT 6
PCT-US04-36404-145
; Sequence 145, Application PC/TUS0436404
; GENERAL INFORMATION:
; APPLICANT: Mayo Foundation for Medical Education and Research
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Burgess, Christopher
; APPLICANT: Myerow, Susan
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Maimonis, Peter
; APPLICANT: Molino, Gary
; APPLICANT: Burgart, Lawrence
; APPLICANT: Boardman, Lisa A.
; APPLICANT: Thibodeau, Steven
; APPLICANT: Lewis, Marcia
; TITLE OF INVENTION: Use of Differentially Expressed Nucleic Acid Sequences as
; TITLE OF INVENTION: Biomarkers for Cancer
; FILE REFERENCE: 1657/2022
; CURRENT APPLICATION NUMBER: PCT/US04/36404
; CURRENT FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US04-36404-145

```

[illegible]

RESULT 7  
 US-10-170-205E-19124  
 ; Sequence 19124, Application US/10170205E  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ADAMS, Mark  
 ; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN  
 ; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF  
 ; FILE REFERENCE: CLO01381  
 ; CURRENT APPLICATION NUMBER: US/10/170,205E  
 ; CURRENT FILING DATE: 2002-06-13  
 ; NUMBER OF SEQ ID NOS: 40312  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 19124  
 ; LENGTH: 112  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-170-205E-19124

|                       |                  |                    |           |             |
|-----------------------|------------------|--------------------|-----------|-------------|
| Query Match           | 65.0%            | Score 356;         | DB 27;    | Length 112; |
| Best Local Similarity | 65.7%            | Pred. NO. 3.8e-32; |           |             |
| Matches 69;           | Conservative 14; | Mismatches 22;     | Indels 0; | Gaps 0;     |

|    |    |   |     |
|----|----|---|-----|
| Qy | 2  | SGSOLAAVALLILLVLSQAQGVVIKYHGFQVQLSEYKVLNELEEKQMSDPQOQKSGLLPDV | 61  |
| Db | 7  | SGLLPGVAVLLLLQLSTQSVYIQGFRVQLSEMKLSLDLEAQWAPSPRIQAOSLLPAP     | 66  |
| Qy | 62 | CYNPALPDLQLPVCASQEAASFTKALRTIATDECELCINVACTGC                 | 106 |
| Db | 67 | CHHPALPDOLPVCASQEAASSFKTLRTIANDCELQVNACTGC                    | 111 |

```

RESULT 8
US-10-479-606-5
; Sequence 5, Application US/10479606
; GENERAL INFORMATION:
; APPLICANT: Cetin, Yalcin
; APPLICANT: Savas, Yukse1
; TITLE OF INVENTION: Guanylate-cyclase C 1
; TITLE OF INVENTION: treatment of respira
; FILE REFERENCE: 03100192aa
; CURRENT APPLICATION NUMBER: US/10/479,606
; CURRENT FILING DATE: 2003-12-04
; PRIOR APPLICATION NUMBER: DE10127119.0
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/DE02/02040
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 112
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-479-606-5

```

|       |                       |   |                    |                |             |         |
|-------|-----------------------|---|--------------------|----------------|-------------|---------|
|       | Query Match           | 65.0%   | Score 356;         | DB 30;         | Length 112; |         |
|       | Best Local Similarity | 65.7%;  | Pred. No. 3.8e-32; |                |             |         |
|       | Matches               | 69;   | Conservative 14;   | Mismatches 22; | Indels 0;   | Gaps 0; |
| <hr/> |                       |   |                    |                |             |         |
| Qy    | 2                     | SGSQLWAVLLLVLSAQGVYIKYGFGVQLVESVKLMLEBKQMSPDQQKXGSLLPDV   | 61                 |                |             |         |
|       | :                     | : :   | :                  | :              | :           | :       |
|       | :                     | : :   | :                  | :              | :           | :       |
| Dd    | 7                     | SGLLPGVAVVLLLLSTQSVIYIQGFVRQLSEMKKLSDLEAQAAPSRLRQAQSLIPAV | 66                 |                |             |         |
|       | :                     | : :   | :                  | :              | :           | :       |
|       | :                     | : :   | :                  | :              | :           | :       |
| <hr/> |                       |   |                    |                |             |         |
| Qy    | 62                    | CYNPALPLDIOPVCASQEAAATFKALRTTATDECELGINVACTGC             | 106                |                |             |         |
|       | :                     | : :   | :                  | :              | :           | :       |
|       | :                     | : :   | :                  | :              | :           | :       |
| <hr/> |                       |   |                    |                |             |         |
| Dd    | 67                    | CHHPALPDLIQPVCASGEASSIFKTITTIANDDCELCVNAVCTGC             | 111                |                |             |         |
|       | :                     | : :   | :                  | :              | :           | :       |
|       | :                     | : :   | :                  | :              | :           | :       |

```

RESULT 9
US-10-700-439-145
; Sequence 145, Application US/10700439
; GENERAL INFORMATION:
; APPLICANT: Mayo Foundation for Medical Education and Research
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Burgess, Christopher
; APPLICANT: Myerow, Susan
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Maimonis, Peter
; APPLICANT: Molino, Gary
; APPLICANT: Burgart, Lawrence
; APPLICANT: Boardman, Lisa A.
; APPLICANT: Thibodeau, Steven
; APPLICANT: Lewis, Marcia
; TITLE OF INVENTION: Use of Differentially Expressed Nucleic Acids to Diagnose Cancer
; FILE REFERENCE: 1657/2022
; CURRENT APPLICATION NUMBER: US/10700.439
; CURRENT FILING DATE: 2003-11-04
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 145
; LENGTH: 112
; TYPE: prt
; ORGANISM: Homo sapiens
US-10-700-439-145

```





```
RESULT 14
US-60-568-073-1020
; Sequence 1020, Application US/60568073
; GENERAL INFORMATION:
; APPLICANT: Huang, Fei
; APPLICANT: Han, Xia
; APPLICANT: Shaw, Peter
; APPLICANT: Clark, Edwin
; TITLE OF INVENTION: COLON CANCER BIOMARKERS
; FILE REFERENCE: 10205 PSP
; CURRENT APPLICATION NUMBER: US/60/568,073
; CURRENT FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 1305
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 1020
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-568-073-1020
Query Match 65.0%; Score 356; DB 37; Length 112;
Best Local Similarity 65.7%; Pred. No. 3.8e-32;
Matches 69; Conservative 14; Mismatches 22; Indels 0; Gaps 0;
Qy 2 SCSQLMAVLLLLVLQSAQGVYIKTHGFQVQLESVKLNLEEEKQMSDPQOQKSGLLPDV 61
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7 SGLLPGVAVLLLLQLQSTQSVYIQGFRVQLESMMKLSLEAQWAPSPRLQAQSLPAV 66
Qy 62 CYNPALPLDLPVCASQEAASSTFKALRTIATDECELCINVACTGC 106
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
67 CHHPALPDQLPVCASQEAASSIFKTLRTIANDDCELCVNVACTGC 111

RESULT 15
US-09-834-366-26398
; Sequence 26398, Application US/09834366
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81.US2.REG
; CURRENT APPLICATION NUMBER: US/09/834,366
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/197,873
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 26398
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-366-26398
Query Match 55.8%; Score 306; DB 23; Length 93;
Best Local Similarity 67.5%; Pred. No. 1.7e-26;
Matches 56; Conservative 12; Mismatches 15; Indels 0; Gaps 0;
Qy 24 IKYHGFQVLESVKLNLEEEKQMSDPQOQKSGLLPDVCYNPALPLDLPVCASQEAAS 83
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
10 LKYQGFRVQLESMMKLSLEAQWAPSPRLQAQSLPAVCHHPALPDQLPVCASQEAASSI 69
Qy 84 FKALRTIATDECELCINVACTGC 106
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
70 FKLRTIANDDCELCVNVACTGC 92
```

Search completed: March 26, 2005, 17:10:17  
Job time : 437.927 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2005, 16:44:35 ; Search time 32.8832 Seconds  
(without alignments)  
187.395 Million cell updates/sec

Title: US-10-775-481A-55  
Perfect score: 548  
Sequence: 1 MSGSOLMAVLLLVLSAQ.....LRTIATDCELCINVACTGC 106

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 260697 seqs, 58133403 residues

Total number of hits satisfying chosen parameters: 260697

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents\_AA\_New.\*  
1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US11\_NEW\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description          |
|------------|-------|-------------|--------|-------|----------------------|
| 1          | 81    | 14.8        | 16     | 7     | US-11-066-697-1249   |
| 2          | 73    | 13.3        | 1139   | 1     | PCT-US05-03483-6     |
| 3          | 69.5  | 12.7        | 1959   | 1     | PCT-US04-42360-1382  |
| 4          | 69.5  | 12.7        | 1959   | 1     | PCT-US04-42360-1384  |
| 5          | 68    | 12.4        | 201    | 8     | US-60-655-875-169824 |
| 6          | 68    | 12.4        | 295    | 8     | US-60-655-875-162848 |
| 7          | 66.5  | 12.1        | 212    | 8     | US-10-450-763-34297  |
| 8          | 65.5  | 12.0        | 218    | 8     | US-60-655-875-161554 |
| 9          | 65.5  | 12.0        | 218    | 8     | US-60-655-875-161177 |
| 10         | 65.5  | 12.0        | 271    | 8     | US-60-655-875-162900 |
| 11         | 65.5  | 12.0        | 516    | 6     | US-10-483-448-1225   |
| 12         | 65    | 11.9        | 341    | 6     | US-10-450-763-36316  |
| 13         | 64.5  | 11.8        | 1225   | 6     | US-10-450-763-41984  |
| 14         | 64.5  | 11.8        | 1307   | 6     | US-10-450-763-41544  |
| 15         | 64    | 11.7        | 331    | 8     | US-60-643-717-2766   |
| 16         | 64    | 11.7        | 700    | 7     | US-11-031-175-14225  |
| 17         | 63    | 11.5        | 304    | 6     | US-10-450-763-37880  |
| 18         | 63    | 11.5        | 539    | 6     | US-10-450-763-36871  |
| 19         | 63    | 11.5        | 539    | 6     | US-10-450-763-56899  |
| 20         | 63    | 11.5        | 840    | 1     | PCT-US04-32678-1     |
| 21         | 63    | 11.5        | 840    | 6     | US-10-679-102-1      |
| 22         | 63    | 11.5        | 840    | 6     | US-10-436-715A-39    |
| 23         | 63    | 11.5        | 840    | 6     | US-10-436-715A-69    |
| 24         | 62.5  | 11.4        | 166    | 8     | US-60-655-875-165482 |
| 25         | 62.5  | 11.4        | 321    | 6     | US-10-450-763-57226  |

ALIGNMENTS

RESULT 1

US-11-066-697-1249  
; Sequence 1249, Application US/11066697  
; GENERAL INFORMATION:  
; APPLICANT: Bridon, Dominique P.  
; APPLICANT: Ezrin, Alan M.  
; APPLICANT: Milner, Peter G.  
; APPLICANT: Holmes, Darren L.  
; APPLICANT: Thibaudau, Karen  
; TITLE OF INVENTION: PROTECTION OF ENDOGENOUS THERAPEUTIC PEPTIDES FROM  
; TITLE OF INVENTION: PEPTIDASE ACTIVITY THROUGH CONJUGATION TO BLOOD  
; TITLE OF INVENTION: COMPONENTS  
; FILE REFERENCE: 500862002301  
; CURRENT APPLICATION NUMBER: US/11/066,697  
; CURRENT FILING DATE: 2005-02-25  
; PRIOR APPLICATION NUMBER: 09/657,276  
; PRIOR FILING DATE: 2000-09-07  
; PRIOR APPLICATION NUMBER: 60/153,406  
; PRIOR FILING DATE: 1999-09-10  
; PRIOR APPLICATION NUMBER: 60/159,783  
; PRIOR FILING DATE: 1999-10-15  
; NUMBER OF SEQ ID NOS: 1617  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1249  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-11-066-697-1249

Query Match 14.8%; Score 81; DB 7; Length 16;  
Best Local Similarity 85.7%; Pred. No. 0.0043;  
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 93 DECELCINVACTGC 106  
|:||||:|||||  
Db 2 DDCELCNVACTGC 15

RESULT 2

PCT-US05-03483-6  
; Sequence 6, Application PC/TUS0503483  
; GENERAL INFORMATION:  
; APPLICANT: EXELIXIS, INC.  
; TITLE OF INVENTION: MAN2AS AS MODIFIERS OF THE IGFR PATHWAY AND METHODS OF USE  
; FILE REFERENCE: EX05-005



```

; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijiang
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/60/655,875
; CURRENT FILING DATE: 2005-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 162848
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATURE:
; OTHER INFORMATION: Homolog annotation: Hit ID=NP_506153.1; Match level="QueryCoverage
; OTHER INFORMATION: =99%, HitCoverage=15%, E-value=3e-53, Identity=38%"; Hit descrip
; OTHER INFORMATION: =TEMO (S987) [Caenorhabditis elegans] emb|CAA98425.1| Hypothesica
; OTHER INFORMATION: protein C14C10.5 [Caenorhabditis elegans
; FEATURE:
; OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID_89271; Strand=+; Position=2-
US-60-655-875-162848

Query Match 12.4%; Score 68; DB 8; Length 295;
Best Local Similarity 23.7%; Pred. No. 5.1;
Matches 27; Conservative 17; Mismatches 38; Indels 32; Gaps 4;

Qy 5 QLVAAVLLLVLSAQG-----VYIKYHGFQVLESVKKLNLEBKQMSDPQQQ 53
Db 45 QLVAAVLLLVLSAQG-----VYIKYHGFQVLESVKKLNLEBKQMSDPQQQ 53

Qy 54 KSGLLPDVCYNPALPL-DLPVCASQEAATF-----KALRTIATDE 94
Db 101 ----YPGICHKPAWPLDPKQKLGFEAREICGNRENFLALAEKLLTIKDE 150

RESULT 7
US-10-450-763-34297
; Sequence 34297, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 34297
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-450-763-34297

Query Match 12.1%; Score 66.5; DB 6; Length 218;
Best Local Similarity 32.6%; Pred. No. 5.2;
Matches 30; Conservative 15; Mismatches 32; Indels 15; Gaps 5;

Qy 1 MSGQLAAVLLLVLSAQGQVYIKYHGFQVLESVKKLNLEBKQMSDPQQQKSGLLPD 60
Db 107 MERNQLAAVLLLVLSAQGQVYIKYHGFQVLESVKKLNLEBKQMSDPQQQKSGLLPD 158
```

```

Qy 61 VCVNPALPLDLQ-PVCAS-QEASTFKALRTI 90
Db 159 -----ALGSALQGAQVSWPQLLSTYRLHSL 185

RESULT 8
US-60-655-875-161554
; Sequence 161554, Application US/60655875
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijiang
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/60/655,875
; CURRENT FILING DATE: 2005-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 161554
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATURE:
; OTHER INFORMATION: Homolog annotation: Hit ID=CAE60769.1; Match level="QueryCoverage
; OTHER INFORMATION: =100%, HitCoverage=66%, E-value=6e-80, Identity=64%"; Hit descrip
; OTHER INFORMATION: Hypothetical protein CBG04457 [Caenorhabditis briggsae]
; FEATURE:
; OTHER INFORMATION: Gene Ontology: Mol. function=protein binding; Cellular component
; OTHER INFORMATION: =cytoplasm; Biological process=protein folding
; OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID_87977; Strand=+; Position=3-
US-60-655-875-161554

Query Match 12.0%; Score 65.5; DB 8; Length 212;
Best Local Similarity 27.2%; Pred. No. 6.6;
Matches 22; Conservative 10; Mismatches 34; Indels 15; Gaps 3;

Qy 34 ESVKKLNLEBKQMSDP-----QQQKSGLLPDVC--YNPALPLDLQPVCSQEA 80
Db 78 ELAKEVKEMEQKAYIDPELSDEKKNLGNQLFKKGDYFGAMRHYNEAIKRNPDNAVLYSNR 137

Qy 81 ASTFKALRTI--ATDECELGI 99
Db 138 AACFTKLMEFQRAVDDCDLCL 158

RESULT 9
US-60-655-875-161177
; Sequence 161177, Application US/60655875
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijiang
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/60/655,875
; CURRENT FILING DATE: 2005-02-24
```

```

; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 161177
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATURE:
; OTHER INFORMATION: Homolog annotation: Hit ID=CAB60769.1; Match level="QueryCoverage
; OTHER INFORMATION: =100%, HitCoverage=68%, E-value=5e-83, Identity=64%"; Hit descrip
; OTHER INFORMATION: Hypothetical protein CBG04457 [Caenorhabditis briggsae]
; FEATURE:
; OTHER INFORMATION: Gene Ontology: Mol. function=protein binding; Cellular component
; OTHER INFORMATION: =cytoplasm; Biological process=protein folding
; FEATURE:
; OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID_87600; Strand=+; Position=1-
US-60-655-875-161177

Query Match      12.0%; Score 65.5; DB 8; Length 218;
Best Local Similarity 27.2%; Pred. No. 6.8;
Matches 22; Conservative 10; Mismatches 34; Indels 15; Gaps 3;

QY 34 ESVKKLELEEKQMSDP-----QQQKSGLLPDVC--YNPALPLDLQPVCAQEA 80
DB 83 ELAKEVKEMEQKAYIDPELSDKEKNLGNQLFKKGYDPCGAMRHYNEAIKRNPNNAVLYSNR 142
QY 81 ASTFKALRTI--ATDECELICI 99
DB 143 AACFTKLMFQRAVDDCDLCL 163

RESULT 10
US-60-655-875-162900
; Sequence 162900, Application US/60655875
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; FILE REFERENCE: IN PLANTS AND COMPOSITIONS THEREOF
; CURRENT APPLICATION NUMBER: US/60/655,875
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 162900
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATURE:
; OTHER INFORMATION: Homolog annotation: Hit ID=CAB60769.1; Match level="QueryCoverage
; OTHER INFORMATION: =99%, HitCoverage=84%, E-value=1e-101, Identity=62%"; Hit descrip
; OTHER INFORMATION: Hypothetical protein CBG04457 [Caenorhabditis briggsae]
; FEATURE:
; OTHER INFORMATION: Gene Ontology: Mol. function=UDP-N-acetylglucosamine-peptide N
; OTHER INFORMATION: =acetylglucosaminyltransferase activity; Cellular component
; OTHER INFORMATION: =perinuclear space; Biological process=protein folding
; FEATURE:
; OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID_89323; Strand=+; Position=84-
US-60-655-875-162900

Query Match      12.0%; Score 65.5; DB 8; Length 271;
Best Local Similarity 27.2%; Pred. No. 9;
Matches 22; Conservative 10; Mismatches 34; Indels 15; Gaps 3;

QY 34 ESVKKLELEEKQMSDP-----QQQKSGLLPDVC--YNPALPLDLQPVCAQEA 80
DB 122 ELAKEVKEMEQKAYIDPELSDKEKNLGNQLFKKGYDPCGAMRHYNEAIKRNPNNAVLYSNR 181
```

```

QY 81 ASTFKALRTI--ATDECELICI 99
DB 182 AACFTKLMFQRAVDDCDLCL 202

RESULT 11
US-10-489-448-1225
; Sequence 1225, Application US/10489448
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wang, Dunrui
; APPLICANT: Ma, Yungqing
; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Gezhi
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 810CIP PCT
; CURRENT APPLICATION NUMBER: US/10/489,448
; CURRENT FILING DATE: 1004-03-10
; PRIOR APPLICATION NUMBER: US 60/324,631
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: PCT/US00/35017
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: PCT/US01/03800
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 3476
; SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 1225
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-489-448-1225

Query Match      12.0%; Score 65.5; DB 6; Length 516;
Best Local Similarity 31.6%; Pred. No. 20;
Matches 25; Conservative 15; Mismatches 26; Indels 13; Gaps 4;

QY 10 VLLLLVLSAQG-----VVKYHGFQVQLSVKLNLELEKQMSDPQQKSGLLPDVC 62
DB 71 LLELLVLSQFLSILPEELQIWVQQHNPSSGEEAVTLLEDL-ERFDDPQQQ-----VPASP 125
QY 63 YNPALP-LDLOQPVCAQEA 80
DB 126 QGPAVPWKDLICLRASQES 144

RESULT 12
```

```
US-10-450-763-36316
; Sequence 36316, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 36316
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (91)..(204)
; OTHER INFORMATION: ZAP domain identified by Pfam, accession name ZAP, E-value=
; OTHER INFORMATION: 6.3e-19, Pfam score of 76.3
US-10-450-763-36316
Query Match 11.9%; Score 65; DB 6; Length 341;
Best Local Similarity 22.7%; Pred. No. 14;
Matches 17; Conservative 21; Mismatches 21; Indels 16; Gaps 3;
Qy 23 YIKYHGFQVLESVKKLENE-----LEEKQMSDPQKSGLLPDVCYNPALPDLQPVCA 76
Db 156 YVDYKQYQYDI-TVSDLNQPMVLVLLKKRNDNSEAQLAHLPCLCFTGL----- 205
Qy 77 SQEAASTFKALRTIA 91
Db 206 TDQATSDPQLMKAVA 220
RESULT 13
US-10-450-763-41984
; Sequence 41984, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 41984
; LENGTH: 1225
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (863)..(908)
; OTHER INFORMATION: kw TRANSCRIPTASE REVERSE II ORF2 domain identified by
; OTHER INFORMATION: eMATRIX, accession number DM01354N, p-value=1.000e-40, raw score
; OTHER INFORMATION: 13.17
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1225)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-41984
```

```
US-10-450-763-41544
; Sequence 41544, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 41544
; LENGTH: 1307
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (896)..(941)
; OTHER INFORMATION: kw TRANSCRIPTASE REVERSE II ORF2 domain identified by
; OTHER INFORMATION: eMATRIX, accession number DM01354N, p-value=1.000e-40, raw score
; OTHER INFORMATION: 13.17
US-10-450-763-41544
Query Match 11.8%; Score 64.5; DB 6; Length 1307;
Best Local Similarity 22.8%; Pred. No. 83;
Matches 26; Conservative 18; Mismatches 37; Indels 33; Gaps 4;
Qy 2 SGSOLMAA-----VLLLLVLSAQAGVYIKYHGFQVLESVKKLELEEKQMSDPQK 54
Db 406 SGSNICCAPISAVLQQLLLIPROTGSVDLQQTPTDLQLR-VLVNVRRTNKKQKQHPHQ-- 462
Qy 55 SGLLPDVCYNPALPDLQPVCAQEAASATFKA-----LRTIATDECELCINVA 102
Db 463 -----NPICMSPSSKTKESLKDLMELKTKARELDCDECTSL 499
RESULT 15
US-60-643-717-2766
; Sequence 2766, Application US/60643717
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)A
; CURRENT APPLICATION NUMBER: US/60/643,717
; CURRENT FILING DATE: 2005-01-12
; NUMBER OF SEQ ID NOS: 19247
; SEQ ID NO 2766
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Crocosphaera watsonii WH 8501
US-60-643-717-2766
Query Match 11.7%; Score 64; DB 8; Length 331;
Best Local Similarity 25.0%; Pred. No. 17;
```

|    | Matches | 21; Conservative   | 14; Mismatches | 35; Indels  | 14; Gaps | 4; |
|----|---------|--|----------------|-------------|----------|----|
| Qy | 17      | QSAQGVYIKHGFQVQLSVKKNLEBEKQMSDPQQKSGLLPDCVYCNPALPDLQFVCA     | 76             |             |          |    |
|    |         | : : : : : :  | : : : : : :    | : : : : : : |          |    |
| Db | 146     | QRQKGVY-----DASIRAEKLNQLGYGH--DP-----NLVLDLVYNNPPISPN-KKFSLS | 191            |             |          |    |
| Qy | 77      | SQEAASFTFKALRTIATDECELSGIN                                   | 100            |             |          |    |
|    |         | : : : : : :  | : : : : : :    | : : : : : : |          |    |
| Db | 192     | TPEQIKLEKDYKGNFLKDNFNICFN                                    | 215            |             |          |    |

Search completed: March 26, 2005, 17:11:53  
Job time : 32.8832 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model  
Run on: March 26, 2005, 16:39:19 ; Search time 462.715 Seconds  
(without alignments)  
282.715 Million cell updates/sec

Title: US-10-775-481A-56  
Perfect score: 579  
Sequence: 1 MGRASGLPGVAVLLLL.....RTIANDCELNVNACTGCL 112

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 6959266 seqs, 1168006243 residues

Total number of hits satisfying chosen parameters: 6959266

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents\_AA\_Main:\*  
1: /cgn2\_6/ptodata/1/paa/pctus\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US080 COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US097A\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US097B\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US099A\_COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US099B\_COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*  
28: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*  
29: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep.\*  
30: /cgn2\_6/ptodata/1/paa/US104\_COMB.pep.\*  
31: /cgn2\_6/ptodata/1/paa/US105\_COMB.pep.\*  
32: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep.\*  
33: /cgn2\_6/ptodata/1/paa/US107\_COMB.pep.\*  
34: /cgn2\_6/ptodata/1/paa/US108\_COMB.pep.\*  
35: /cgn2\_6/ptodata/1/paa/US109\_COMB.pep.\*  
36: /cgn2\_6/ptodata/1/paa/US110\_COMB.pep.\*  
37: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description                               |
|------------|-------|---------------|--------|-------|---|
| 1          | 579   | 100.0         | 112    | 1     | PCT-US04-03765-56<br>Sequence 56, Appl    |
| 2          | 579   | 100.0         | 112    | 1     | PCT-US04-36404-145<br>Sequence 145, App   |
| 3          | 579   | 100.0         | 112    | 27    | US-10-170-205B-19124<br>Sequence 19124, A |
| 4          | 579   | 100.0         | 112    | 30    | US-10-479-606-5<br>Sequence 5, Appl       |
| 5          | 579   | 100.0         | 112    | 33    | US-10-700-439-145<br>Sequence 145, App    |
| 6          | 579   | 100.0         | 112    | 33    | US-10-775-481A-56<br>Sequence 56, Appl    |
| 7          | 579   | 100.0         | 112    | 37    | US-60-453-050-12227<br>Sequence 12227, A  |
| 8          | 579   | 100.0         | 112    | 37    | US-60-453-135-12227<br>Sequence 12227, A  |
| 9          | 579   | 100.0         | 112    | 37    | US-60-466-412-12227<br>Sequence 12227, A  |
| 10         | 579   | 100.0         | 112    | 37    | US-60-568-073-1020<br>Sequence 1020, Ap   |
| 11         | 571   | 98.6          | 122    | 37    | US-60-245-228-223<br>Sequence 223, App    |
| 12         | 567   | 97.9          | 111    | 37    | US-60-248-592-71<br>Sequence 71, Appl     |
| 13         | 440   | 76.0          | 93     | 23    | US-09-834-366-26398<br>Sequence 26398, A  |
| 14         | 440   | 76.0          | 93     | 37    | US-60-197-873-26398<br>Sequence 26398, A  |
| 15         | 356   | 61.5          | 106    | 1     | PCT-US04-03765-55<br>Sequence 55, Appl    |
| 16         | 356   | 61.5          | 106    | 33    | US-10-775-481A-55<br>Sequence 150, App    |
| 17         | 271.5 | 46.9          | 89     | 37    | US-60-177-670-190<br>Sequence 6, Appl     |
| 18         | 252   | 43.5          | 109    | 30    | US-10-479-606-6<br>Sequence 1527, Ap      |
| 19         | 216   | 37.3          | 85     | 1     | PCT-US00-05883-1527<br>Sequence 1527, Ap  |
| 20         | 216   | 37.3          | 85     | 24    | US-09-525-299-1527<br>Sequence 68, Appl   |
| 21         | 162.5 | 28.1          | 94     | 1     | PCT-US04-18751-68<br>Sequence 70, Appl    |
| 22         | 162.5 | 28.1          | 115    | 23    | US-09-834-366-17911<br>Sequence 17911, A  |
| 23         | 162.5 | 28.1          | 115    | 25    | US-09-981-353-61<br>Sequence 61, Appl     |
| 24         | 162.5 | 28.1          | 115    | 26    | US-10-003-608-22<br>Sequence 22, Appl     |
| 25         | 162.5 | 28.1          | 115    | 27    | US-10-170-205B-19125<br>Sequence 19125, A |
| 26         | 162.5 | 28.1          | 115    | 28    | US-10-235-994-22<br>Sequence 22, Appl     |
| 27         | 162.5 | 28.1          | 115    | 28    | US-10-262-473-12<br>Sequence 12, Appl     |
| 28         | 162.5 | 28.1          | 115    | 30    | US-10-479-606-4<br>Sequence 4, Appl       |
| 29         | 162.5 | 28.1          | 115    | 37    | US-60-197-873-17911<br>Sequence 17911, A  |
| 30         | 162.5 | 28.1          | 115    | 37    | US-60-239-841-61<br>Sequence 61, Appl     |
| 31         | 162.5 | 28.1          | 115    | 37    | US-60-453-050-12228<br>Sequence 12228, A  |
| 32         | 162.5 | 28.1          | 115    | 37    | US-60-453-135-12228<br>Sequence 12228, A  |
| 33         | 162.5 | 28.1          | 115    | 37    | US-60-466-412-12228<br>Sequence 12228, A  |
| 34         | 162.5 | 28.1          | 108    | 1     | PCT-US04-18751-72<br>Sequence 72, Appl    |
| 35         | 136   | 23.5          | 108    | 28    | US-10-262-473-16<br>Sequence 16, Appl     |
| 36         | 136   | 23.5          | 108    | 28    | US-10-262-473-16<br>Sequence 16, Appl     |
| 37         | 135   | 23.3          | 30     | 37    | US-60-187-385-887<br>Sequence 887, App    |
| 38         | 135   | 23.3          | 30     | 37    | US-60-234-446-896<br>Sequence 896, App    |
| 39         | 134   | 23.1          | 24     | 1     | PCT-US04-18751-77<br>Sequence 77, Appl    |
| 40         | 132   | 22.8          | 102    | 1     | PCT-US04-18751-71<br>Sequence 71, Appl    |
| 41         | 132   | 22.8          | 102    | 28    | US-10-262-473-14<br>Sequence 14, Appl     |
| 42         | 118.5 | 20.5          | 23     | 1     | PCT-US04-18751-78<br>Sequence 78, Appl    |
| 43         | 106   | 18.3          | 24     | 37    | US-60-160-203-4187<br>Sequence 4187, Ap   |
| 44         | 106   | 18.3          | 24     | 37    | US-60-169-840-6139<br>Sequence 6139, Ap   |
| 45         | 100   | 17.3          | 17     | 37    | US-60-160-203-3643<br>Sequence 3643, Ap   |

ALIGNMENTS

RESULT 1  
PCT-US04-03765-56  
; Sequence 56, Application PC/TUS0403765  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; APPLICANT: Pitari, Giovanni Mario  
; APPLICANT: Park, Jason  
; APPLICANT: Schulz, Stephanie  
; APPLICANT: Wolfe, Henry R.  
; APPLICANT: Lubbe, Wilhelm  
; TITLE OF INVENTION: The Use Of GCC Ligands  
; FILE REFERENCE: 08321-168 PC1  
; CURRENT FILING DATE: 2004-02-10  
; PRIOR APPLICATION NUMBER: PCT/US04/03765  
; PRIOR FILING DATE: 2004-02-10  
; PRIOR APPLICATION NUMBER: US 60/446,730  
; PRIOR FILING DATE: 2003-02-10  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 56

```
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US04-03765-56

Query Match      100.0%; Score 579; DB 1; Length 112;
Best Local Similarity 100.0%; Pred. No. 7e-56;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGCRAASGLLPGVAVVLLLLQSTQSVYIQGFRVQLESMMKLSDLAQWAPSPRLQAO 60
    |||||
Db 1 MGCRAASGLLPGVAVVLLLLQSTQSVYIQGFRVQLESMMKLSDLAQWAPSPRLQAO 60
    |||||
QY 61 SLLPAVCHHPALPDQLQPVCSQEQASSIFKTLRTIANDDCELCVNACTGCL 112
    |||||
Db 61 SLLPAVCHHPALPDQLQPVCSQEQASSIFKTLRTIANDDCELCVNACTGCL 112
    |||||

RESULT 2
PCT-US04-36404-145
; Sequence 145, Application PC/TUS0436404
; GENERAL INFORMATION:
; APPLICANT: Mayo Foundation for Medical Education and Research
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Burgess, Christopher
; APPLICANT: Myerow, Susan
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Maimonis, Peter
; APPLICANT: Molino, Gary
; APPLICANT: Burtgart, Lawrence
; APPLICANT: Boardman, Lisa A.
; APPLICANT: Thibodeau, Steven
; APPLICANT: Lewis, Marcia
; TITLE OF INVENTION: Use of Differentially Expressed Nucleic Acid Sequences as
; FILE REFERENCE: 1657/2022
; CURRENT APPLICATION NUMBER: PCT/US04/36404
; CURRENT FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US04-36404-145

Query Match      100.0%; Score 579; DB 1; Length 112;
Best Local Similarity 100.0%; Pred. No. 7e-56;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGCRAASGLLPGVAVVLLLLQSTQSVYIQGFRVQLESMMKLSDLAQWAPSPRLQAO 60
    |||||
Db 1 MGCRAASGLLPGVAVVLLLLQSTQSVYIQGFRVQLESMMKLSDLAQWAPSPRLQAO 60
    |||||
QY 61 SLLPAVCHHPALPDQLQPVCSQEQASSIFKTLRTIANDDCELCVNACTGCL 112
    |||||
Db 61 SLLPAVCHHPALPDQLQPVCSQEQASSIFKTLRTIANDDCELCVNACTGCL 112
    |||||

RESULT 3
US-10-170-205E-19124
; Sequence 19124, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19124
; LENGTH: 112
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-19124

Query Match      100.0%; Score 579; DB 27; Length 112;
Best Local Similarity 100.0%; Pred. No. 7e-56;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGCRAASGLLPGVAVVLLLLQSTQSVYIQGFRVQLESMMKLSDLAQWAPSPRLQAO 60
    |||||
Db 1 MGCRAASGLLPGVAVVLLLLQSTQSVYIQGFRVQLESMMKLSDLAQWAPSPRLQAO 60
    |||||
QY 61 SLLPAVCHHPALPDQLQPVCSQEQASSIFKTLRTIANDDCELCVNACTGCL 112
    |||||
Db 61 SLLPAVCHHPALPDQLQPVCSQEQASSIFKTLRTIANDDCELCVNACTGCL 112
    |||||

RESULT 4
US-10-479-606-5
; Sequence 5, Application US/10479606
; GENERAL INFORMATION:
; APPLICANT: Cetin, Yalcin
; APPLICANT: Savas, Yusef
; TITLE OF INVENTION: Guanylate-cyclase C ligand, administered via the airways, for the
; FILE REFERENCE: 03100192aa
; CURRENT APPLICATION NUMBER: US/10/479,606
; CURRENT FILING DATE: 2003-12-04
; PRIOR APPLICATION NUMBER: DE10127119.0
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/DE02/02040
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 112
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-479-606-5

Query Match      100.0%; Score 579; DB 30; Length 112;
Best Local Similarity 100.0%; Pred. No. 7e-56;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGCRAASGLLPGVAVVLLLLQSTQSVYIQGFRVQLESMMKLSDLAQWAPSPRLQAO 60
    |||||
Db 1 MGCRAASGLLPGVAVVLLLLQSTQSVYIQGFRVQLESMMKLSDLAQWAPSPRLQAO 60
    |||||
QY 61 SLLPAVCHHPALPDQLQPVCSQEQASSIFKTLRTIANDDCELCVNACTGCL 112
    |||||
Db 61 SLLPAVCHHPALPDQLQPVCSQEQASSIFKTLRTIANDDCELCVNACTGCL 112
    |||||

RESULT 5
US-10-700-439-145
; Sequence 145, Application US/10700439
; GENERAL INFORMATION:
; APPLICANT: Mayo Foundation for Medical Education and Research
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Burgess, Christopher
; APPLICANT: Myerow, Susan
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Maimonis, Peter
; APPLICANT: Molino, Gary
; APPLICANT: Burtgart, Lawrence
; APPLICANT: Boardman, Lisa A.
; APPLICANT: Thibodeau, Steven
; APPLICANT: Lewis, Marcia
; TITLE OF INVENTION: Use of Differentially Expressed Nucleic Acid Sequences as
; FILE REFERENCE: 1657/2022
; CURRENT APPLICATION NUMBER: US/10/700,439
; CURRENT FILING DATE: 2003-11-04
```

```
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-700-439-145

Query Match      100.0%; Score 579; DB 33; Length 112;
Best Local Similarity 100.0%; Pred. No. 7e-56;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGCRAASGLLPQVAVVLLLLQSTQSVYIQGFRVQLESMMKLSDLQAQWAPSRLQAQ 60
Db 1 MGCRAASGLLPQVAVVLLLLQSTQSVYIQGFRVQLESMMKLSDLQAQWAPSRLQAQ 60

Qy 61 SLLPAVCHHPALPDQLPVCASQEQASSIFKTLRTIANDDCELCVNVACTGCL 112
Db 61 SLLPAVCHHPALPDQLPVCASQEQASSIFKTLRTIANDDCELCVNVACTGCL 112

RESULT 6
US-10-775-481A-56
; Sequence 56, Application US/10775481A
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; APPLICANT: Pitari, Giovanni Mario
; APPLICANT: Park, Jason
; APPLICANT: Schulz, Stephanie
; APPLICANT: Wolfe, Henry R.
; APPLICANT: Lubbe, Wilhelm
; TITLE OF INVENTION: The Use Of GCC Ligands
; FILE REFERENCE: 08321-0168 US1
; CURRENT APPLICATION NUMBER: US/10/775,481A
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/446,730
; PRIOR FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-481A-56

Query Match      100.0%; Score 579; DB 33; Length 112;
Best Local Similarity 100.0%; Pred. No. 7e-56;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGCRAASGLLPQVAVVLLLLQSTQSVYIQGFRVQLESMMKLSDLQAQWAPSRLQAQ 60
Db 1 MGCRAASGLLPQVAVVLLLLQSTQSVYIQGFRVQLESMMKLSDLQAQWAPSRLQAQ 60

Qy 61 SLLPAVCHHPALPDQLPVCASQEQASSIFKTLRTIANDDCELCVNVACTGCL 112
Db 61 SLLPAVCHHPALPDQLPVCASQEQASSIFKTLRTIANDDCELCVNVACTGCL 112

RESULT 7
US-60-453-050-12227
; Sequence 12227, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12227
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-12227

Query Match      100.0%; Score 579; DB 37; Length 112;
Best Local Similarity 100.0%; Pred. No. 7e-56;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGCRAASGLLPQVAVVLLLLQSTQSVYIQGFRVQLESMMKLSDLQAQWAPSRLQAQ 60
Db 1 MGCRAASGLLPQVAVVLLLLQSTQSVYIQGFRVQLESMMKLSDLQAQWAPSRLQAQ 60

Qy 61 SLLPAVCHHPALPDQLPVCASQEQASSIFKTLRTIANDDCELCVNVACTGCL 112
Db 61 SLLPAVCHHPALPDQLPVCASQEQASSIFKTLRTIANDDCELCVNVACTGCL 112

RESULT 8
US-60-453-135-12227
; Sequence 12227, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12227
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-12227

Query Match      100.0%; Score 579; DB 37; Length 112;
Best Local Similarity 100.0%; Pred. No. 7e-56;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGCRAASGLLPQVAVVLLLLQSTQSVYIQGFRVQLESMMKLSDLQAQWAPSRLQAQ 60
Db 1 MGCRAASGLLPQVAVVLLLLQSTQSVYIQGFRVQLESMMKLSDLQAQWAPSRLQAQ 60

Qy 61 SLLPAVCHHPALPDQLPVCASQEQASSIFKTLRTIANDDCELCVNVACTGCL 112
Db 61 SLLPAVCHHPALPDQLPVCASQEQASSIFKTLRTIANDDCELCVNVACTGCL 112

RESULT 9
US-60-466-412-12227
; Sequence 12227, Application US/60466412
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001466
; CURRENT APPLICATION NUMBER: US/60/466,412
; CURRENT FILING DATE: 2003-04-30
; NUMBER OF SEQ ID NOS: 429241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12227
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-466-412-12227

Query Match      100.0%; Score 579; DB 37; Length 112;
Best Local Similarity 100.0%; Pred. No. 7e-56;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGCRAASGLLPQVAVVLLLLQSTQSVYIQGFRVQLESMMKLSDLQAQWAPSRLQAQ 60
Db 1 MGCRAASGLLPQVAVVLLLLQSTQSVYIQGFRVQLESMMKLSDLQAQWAPSRLQAQ 60

Qy 61 SLLPAVCHHPALPDQLPVCASQEQASSIFKTLRTIANDDCELCVNVACTGCL 112
Db 61 SLLPAVCHHPALPDQLPVCASQEQASSIFKTLRTIANDDCELCVNVACTGCL 112
```

Db 1 MGCRAASGLLPGVAVVLLLLQSTQSVYIQGFRVQLESMMKLSLEAQWAPSRLQAQ 60  
QY 61 SLLPACHHPALPDQLQPVCAQSEASSIFKTLRTIANDDCELCVNVACTGCL 112  
Db 61 SLLPACHHPALPDQLQPVCAQSEASSIFKTLRTIANDDCELCVNVACTGCL 112

## RESULT 10

US-60-568-073-1020  
; Sequence 1020, Application US/60568073  
; GENERAL INFORMATION:  
; APPLICANT: Huang, Fei  
; APPLICANT: Han, Xia  
; APPLICANT: Shaw, Peter  
; APPLICANT: Clark, Edwin  
; TITLE OF INVENTION: COLON CANCER BIOMARKERS  
; FILE REFERENCE: 10205 PSP  
; CURRENT APPLICATION NUMBER: US/60/568,073  
; CURRENT FILING DATE: 2004-05-04  
; NUMBER OF SEQ ID NOS: 1305  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 1020  
; LENGTH: 112  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-568-073-1020

Query Match 100.0%; Score 579; DB 37; Length 112;  
Best Local Similarity 100.0%; Pred. No. 7e-56;  
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGCRAASGLLPGVAVVLLLLQSTQSVYIQGFRVQLESMMKLSLEAQWAPSRLQAQ 60  
Db 1 MGCRAASGLLPGVAVVLLLLQSTQSVYIQGFRVQLESMMKLSLEAQWAPSRLQAQ 60  
QY 61 SLLPACHHPALPDQLQPVCAQSEASSIFKTLRTIANDDCELCVNVACTGCL 112  
Db 61 SLLPACHHPALPDQLQPVCAQSEASSIFKTLRTIANDDCELCVNVACTGCL 112

## RESULT 11

US-60-245-228-223  
; Sequence 223, Application US/60245228  
; GENERAL INFORMATION:  
; APPLICANT: Beasley, Ellen  
; TITLE OF INVENTION: ISOLATED HUMAN CYCLASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN CYCLASE PROTEINS, AND USES  
; FILE REFERENCE: CL000878  
; CURRENT APPLICATION NUMBER: US/60/245,228  
; CURRENT FILING DATE: 2000-11-03  
; NUMBER OF SEQ ID NOS: 630  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 223  
; LENGTH: 122  
; TYPE: PRT  
; ORGANISM: HUMAN  
US-60-245-228-223

Query Match 98.6%; Score 571; DB 37; Length 122;  
Best Local Similarity 99.1%; Pred. No. 6e-55;  
Matches 111; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGCRAASGLLPGVAVVLLLLQSTQSVYIQGFRVQLESMMKLSLEAQWAPSRLQAQ 60  
Db 11 MGCRAASGLLPGVAVVLLLLQSTQSVYIQGFRVQLESMMKLSLEAQWAPSRLQAQ 70  
QY 61 SLLPACHHPALPDQLQPVCAQSEASSIFKTLRTIANDDCELCVNVACTGCL 112  
Db 71 SLLPACHHPALPDQLQPVCAQSEASSIFKTLRTIANDDCELCVNVACTGCL 122

## RESULT 12

US-60-248-592-71  
; Sequence 71, Application US/60248592  
; GENERAL INFORMATION:  
; APPLICANT: Beasley, Ellen  
; TITLE OF INVENTION: ISOLATED HUMAN CYCLASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN CYCLASE PROTEINS, AND USES  
; FILE REFERENCE: CL000961  
; CURRENT APPLICATION NUMBER: US/60/248,592  
; CURRENT FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 192  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 71  
; LENGTH: 111  
; TYPE: PRT  
; ORGANISM: HUMAN  
US-60-248-592-71

Query Match 97.9%; Score 567; DB 37; Length 111;  
Best Local Similarity 99.1%; Pred. No. 1.5e-54;  
Matches 110; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGCRAASGLLPGVAVVLLLLQSTQSVYIQGFRVQLESMMKLSLEAQWAPSRLQAQ 60  
Db 1 MGCRAASGLLPGVAVVLLLLQSTQSVYIQGFRVQLESMMKLSLEAQWAPSRLQAQ 60  
QY 61 SLLPACHHPALPDQLQPVCAQSEASSIFKTLRTIANDDCELCVNVACTGCL 111  
Db 61 SLLPACHHPALPDQLQPVCAQSEASSIFKTLRTIANDDCELCVNVACTGCL 111

## RESULT 13

US-09-834-366-26398  
; Sequence 26398, Application US/09834366  
; GENERAL INFORMATION:  
; APPLICANT: Bejanin, Stephane  
; APPLICANT: Tanaka, Hiroaki  
; APPLICANT: Dumas Milne Edwards, Jean Baptiste  
; APPLICANT: Jobert, Severin  
; APPLICANT: Giordano, Jean-Yves  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: 81.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/834,366  
; CURRENT FILING DATE: 2001-04-13  
; PRIOR APPLICATION NUMBER: US 60/197,873  
; PRIOR FILING DATE: 2000-04-18  
; NUMBER OF SEQ ID NOS: 52153  
; SOFTWARE: Patent.pm  
; SEQ ID NO 26398  
; LENGTH: 93  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-834-366-26398

Query Match 76.0%; Score 440; DB 23; Length 93;  
Best Local Similarity 91.2%; Pred. No. 1.8e-40;  
Matches 83; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 22 QSTQSVYIQGFRVQLESMMKLSLEAQWAPSRLQAQSLLPVCHHPALPDQLQPVCA 81  
Db 3 RGVKSKDLKYQGFRVQLESMMKLSLEAQWAPSRLQAQSLLPVCHHPALPDQLQPVCA 62  
QY 82 SQEASSIFKTLRTIANDDCELCVNVACTGCL 112  
Db 63 SQEASSIFKTLRTIANDDCELCVNVACTGCL 93

## RESULT 14

US-60-197-873-26398  
; Sequence 26398, Application US/60197873  
; GENERAL INFORMATION:  
; APPLICANT: Bejanin, Stephane

Search completed: March 26, 2005, 17:10:18  
Job time : 463.715 secs

*This Page Blank (uspio)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: March 26, 2005, 16:44:35 ; Search time 34.7445 Seconds  
(without alignments)  
187.395 Million cell updates/sec

Title: US-10-775-481A-56

Perfect score: 579

Sequence: 1 MGRASGLPGVAVLLLL.....RTIANDCELCVNACTGCL 112

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 260697 seqs, 58133403 residues

Total number of hits satisfying chosen parameters: 260697

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA New.\*

- 1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/1/paa/US11\_NEW\_COMB.pep.\*
- 8: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description          |
|------------|-------|-------------|--------|----|----------------------|
| 1          | 96    | 16.6        | 16     | 7  | US-11-066-697-1249   |
| 2          | 74.5  | 12.9        | 985    | 6  | US-10-450-763-47493  |
| 3          | 74    | 12.8        | 483    | 6  | US-10-450-763-33083  |
| 4          | 73    | 12.6        | 351    | 1  | PCT-US04-30360-157   |
| 5          | 71.5  | 12.3        | 301    | 1  | PCT-US04-31036-29    |
| 6          | 71.5  | 12.3        | 301    | 1  | PCT-US04-31036-31    |
| 7          | 68.5  | 11.8        | 166    | 8  | US-60-655-875-165482 |
| 8          | 68.5  | 11.8        | 1444   | 6  | US-10-450-763-46026  |
| 9          | 68    | 11.7        | 376    | 7  | US-11-031-175-15516  |
| 10         | 68    | 11.7        | 1377   | 1  | PCT-US05-02638-105   |
| 11         | 67.5  | 11.7        | 819    | 6  | US-10-450-763-49219  |
| 12         | 67.5  | 11.7        | 953    | 7  | US-11-031-175-14681  |
| 13         | 67    | 11.6        | 187    | 6  | US-10-755-415-236    |
| 14         | 66.5  | 11.5        | 228    | 7  | US-11-033-545-409    |
| 15         | 66.5  | 11.5        | 393    | 6  | US-10-489-448-1493   |
| 16         | 66.5  | 11.5        | 684    | 6  | US-10-450-763-52765  |
| 17         | 66.5  | 11.5        | 824    | 7  | US-11-033-545-312    |
| 18         | 66.5  | 11.5        | 920    | 6  | US-10-450-763-36280  |
| 19         | 66.5  | 11.5        | 920    | 6  | US-10-450-763-52763  |
| 20         | 66.5  | 11.5        | 982    | 6  | US-10-450-763-31580  |
| 21         | 66.5  | 11.5        | 990    | 6  | US-10-450-763-36281  |
| 22         | 66.5  | 11.5        | 990    | 6  | US-10-450-763-52764  |
| 23         | 66.5  | 11.5        | 996    | 6  | US-10-450-763-36284  |
| 24         | 66.5  | 11.5        | 997    | 6  | US-10-450-763-52766  |
| 25         | 66.5  | 11.5        | 1236   | 6  | US-10-450-763-38524  |

|    |      |      |     |   |                      |                   |
|----|------|------|-----|---|----------------------|-------------------|
| 26 | 66   | 11.4 | 776 | 6 | US-10-450-763-58595  | Sequence 58595, A |
| 27 | 65.5 | 11.3 | 737 | 6 | US-10-450-763-45465  | Sequence 45465, A |
| 28 | 65.5 | 11.3 | 737 | 6 | US-10-489-448-2864   | Sequence 2864, Ap |
| 29 | 64   | 11.1 | 315 | 8 | US-60-655-875-131740 | Sequence 131740   |
| 30 | 64   | 11.1 | 345 | 6 | US-10-491-545A-58    | Sequence 58, Appl |
| 31 | 64   | 11.1 | 976 | 6 | US-10-489-448-1519   | Sequence 1519, Ap |
| 32 | 63.5 | 11.0 | 122 | 6 | US-10-450-763-50350  | Sequence 50350, A |
| 33 | 63.5 | 11.0 | 214 | 6 | US-10-450-763-41255  | Sequence 41255, A |
| 34 | 63.5 | 11.0 | 479 | 1 | PCT-US04-17965-1190  | Sequence 1190, Ap |
| 35 | 63.5 | 11.0 | 479 | 1 | PCT-US04-17965B-1190 | Sequence 1190, Ap |
| 36 | 63   | 10.9 | 269 | 6 | US-10-450-763-43513  | Sequence 43513, A |
| 37 | 63   | 10.9 | 353 | 8 | US-60-643-717-8659   | Sequence 8659, Ap |
| 38 | 63   | 10.9 | 445 | 8 | US-60-655-875-146849 | Sequence 146849   |
| 39 | 63   | 10.9 | 745 | 6 | US-10-450-763-56461  | Sequence 56461, A |
| 40 | 62.5 | 10.8 | 166 | 8 | US-60-655-875-158868 | Sequence 158868   |
| 41 | 62.5 | 10.8 | 608 | 1 | PCT-US05-01983-16    | Sequence 16, Appl |
| 42 | 62.5 | 10.8 | 608 | 1 | PCT-US05-01983-90    | Sequence 90, Appl |
| 43 | 62   | 10.7 | 114 | 6 | US-10-526-324-1284   | Sequence 1284, Ap |
| 44 | 62   | 10.7 | 196 | 8 | US-60-655-875-146820 | Sequence 146820   |
| 45 | 62   | 10.7 | 692 | 6 | US-10-450-763-37251  | Sequence 37251, A |

## ALIGNMENTS

## RESULT 1

US-11-066-697-1249

; Sequence 1249, Application US/11066697

; GENERAL INFORMATION: Dominique P.

; APPLICANT: Bridon, Alan M.

; APPLICANT: Ezrin, Alan M.

; APPLICANT: Milner, Peter G.

; APPLICANT: Holmes, Darren L.

; APPLICANT: Thibaudau, Karen

; TITLE OF INVENTION: PROTECTION OF ENDOGENOUS THERAPEUTIC PEPTIDES FROM

; TITLE OF INVENTION: PEPTIDASE ACTIVITY THROUGH CONJUGATION TO BLOOD

; TITLE OF INVENTION: COMPONENTS

; FILE REFERENCE: 500862002301

; CURRENT APPLICATION NUMBER: US/11/066,697

; CURRENT FILING DATE: 2005-02-25

; PRIOR APPLICATION NUMBER: 09/657,276

; PRIOR FILING DATE: 2000-09-07

; PRIOR APPLICATION NUMBER: 60/153,406

; PRIOR FILING DATE: 1999-09-10

; PRIOR APPLICATION NUMBER: 60/159,783

; PRIOR FILING DATE: 1999-10-15

; NUMBER OF SEQ ID NOS: 1617

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 1249

; LENGTH: 16

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Synthetic

; OTHER INFORMATION: Peptide

US-11-066-697-1249

Query Match

Best Local Similarity 16.6%; Score 96; DB 7; Length 16;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 97 NDDCELCVNACTGCL 112

Db 1 NDDCELCVNACTGCL 16

## RESULT 2

US-10-450-763-47493

; Sequence 47493, Application US/10450763

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc

; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

; FILE REFERENCE: 790CIP3/US

CURRENT APPLICATION NUMBER: US/10/450,763  
CURRENT FILING DATE: 2003-06-11  
PRIOR APPLICATION NUMBER: PCT/US01/08631  
PRIOR FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: 09/540,217  
PRIOR FILING DATE: 2000-03-31  
PRIOR APPLICATION NUMBER: 09/649,167  
PRIOR FILING DATE: 2000-08-23  
NUMBER OF SEQ ID NOS: 60736  
SOFTWARE: Custom  
SEQ ID NO 47493  
LENGTH: 985  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (653)...(684)  
OTHER INFORMATION: Protein kinases ATP-binding region proteins domain identified  
OTHER INFORMATION: by eMATRIX, accession number BL00107A, p-value=1.857e-16, raw  
OTHER INFORMATION: score of 18.39  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (521)...(794)  
OTHER INFORMATION: Eukaryotic protein kinase domain identified by Pfam,  
OTHER INFORMATION: accession name pkinase, E-value=1e-68, Pfam score of 241.7  
US-10-450-763-47493

Query Match 12.9%; Score 74.5; DB 6; Length 985;  
Best Local Similarity 24.1%; Pred. No. 6.4;  
Matches 26; Conservative 20; Mismatches 29; Indels 33; Gaps 6;  
QY 12 GVAVVLLLLQSTQSVYIQGFRVQLESN-----KKLSLEAQWAPSPRLQQA 60  
DB 223 GIACARFI---GNRTVIMESLHMQGEIENQITAAFTMIGTSSHLSDKCSQFA----- 271  
QY 61 SLLPVPCHHPALPQDLQPVCAQSEASSIFKLTFTIANDCELCNVAC 108  
DB 272 --IPSLCHYAF-----PYC--DETSSVPKP-RDLCDRDECEILENVLC 308

RESULT 3  
US-10-450-763-33083  
Sequence 33083, Application US/10450763  
GENERAL INFORMATION:  
APPLICANT: Hyseq, Inc  
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
FILE REFERENCE: 790CIP3/US  
CURRENT APPLICATION NUMBER: US/10/450,763  
CURRENT FILING DATE: 2003-06-11  
PRIOR APPLICATION NUMBER: PCT/US01/08631  
PRIOR FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: 09/540,217  
PRIOR FILING DATE: 2000-03-31  
PRIOR APPLICATION NUMBER: 09/649,167  
PRIOR FILING DATE: 2000-08-23  
NUMBER OF SEQ ID NOS: 60736  
SOFTWARE: Custom  
SEQ ID NO 33083  
LENGTH: 483  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-450-763-33083

Query Match 12.8%; Score 74; DB 6; Length 483;  
Best Local Similarity 26.9%; Pred. No. 3.1;  
Matches 28; Conservative 10; Mismatches 36; Indels 30; Gaps 4;  
QY 7 SGLPGVAVVLLLLQSTQSVYIQG---FRVQLESMMKLSLEAQWAP----- 53  
DB 345 SGTVPRI-----LLQNPGRGMDAHSGHTLSFRSQPTRSLSQ-EGLWGPARDRYCHVP 397  
QY 54 -----SPRLQAQSLPVPCHHPALPQDLQPVCAQSEASS 87

DB 398 PLNCPNMHKAPRAVCAVSALRVLCVTHPTSPAGLSFNCSHLQSS 441  
RESULT 4  
PCT-US04-30360-157  
Sequence 157, Application PC/TUS0430360  
GENERAL INFORMATION:  
APPLICANT: PLEXIKON, INC.  
TITLE OF INVENTION: MOLECULAR SCAFFOLDS FOR KINASE LIGAND DEVELOPMENT  
FILE REFERENCE: 039363-1703  
CURRENT APPLICATION NUMBER: PCT/US04/30360  
CURRENT FILING DATE: 2004-09-15  
PRIOR APPLICATION NUMBER: 60/503,277  
PRIOR FILING DATE: 2003-09-15  
NUMBER OF SEQ ID NOS: 167  
SOFTWARE: PatentIn Ver. 3.2  
SEQ ID NO 157  
LENGTH: 351  
TYPE: PRT  
ORGANISM: Unknown Organism  
FEATURE:  
OTHER INFORMATION: Description of Unknown Organism: Mammalian  
OTHER INFORMATION: protein sequence  
PCT-US04-30360-157

Query Match 12.6%; Score 73; DB 1; Length 351;  
Best Local Similarity 32.0%; Pred. No. 2.7;  
Matches 24; Conservative 16; Mismatches 25; Indels 10; Gaps 4;  
QY 41 MKKLSLEAQWAPSPRLQAQSLPVPCHHPAL--PDLQPVCAQSEASSIFKLTFTIAND 98  
DB 274 NRELTEASLERNPVHPRPADLL-----KHEALNPREDQPRCTSLD-SALLERKLLSRK 328  
QY 99 DCELCNVVA---CTG 110  
DB 329 EELEPENIADSSCTG 343

RESULT 5  
PCT-US04-31036-29  
Sequence 29, Application PC/TUS0431036  
GENERAL INFORMATION:  
APPLICANT: Dale Umetsu  
APPLICANT: Rosemarie Dekruyff  
APPLICANT: Jennifer McIntire  
APPLICANT: Gordon Freeman  
TITLE OF INVENTION: T CELL REGULATORY GENES ASSOCIATED WITH  
TITLE OF INVENTION: IMMUNE DISEASE  
FILE REFERENCE: STAN-235CIP  
CURRENT APPLICATION NUMBER: PCT/US04/31036  
CURRENT FILING DATE: 2004-09-15  
PRIOR APPLICATION NUMBER: 60/302,344  
PRIOR FILING DATE: 2001-06-29  
PRIOR APPLICATION NUMBER: 10/188,012  
PRIOR FILING DATE: 2002-07-01  
NUMBER OF SEQ ID NOS: 40  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 29  
LENGTH: 301  
TYPE: PRT  
ORGANISM: H. sapiens  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (1)...(301)  
OTHER INFORMATION: TIM-3, allele 1  
PCT-US04-31036-29

Query Match 12.3%; Score 71.5; DB 1; Length 301;  
Best Local Similarity 28.0%; Pred. No. 3.4;  
Matches 26; Conservative 12; Mismatches 28; Indels 27; Gaps 3;  
QY 10 LFGVAVVLLLLQSTQSVYIQGFRVQLESMMKLSLEAQWAPSPRLQAQSLPVPCHH 69



```
Db 5 LPFDCVLLLLLLLTSRSEVEYRA-----EVGNAYLPCF-YT 41
Qy 70 PALPDQLPVCASQAEASSIFK-----TLRTIAND 98
Db 42 PAAPGNLVPVCGKGCACPFCEGNNVLRTERD 74

RESULT 6
PCT-US04-31036-31
; Sequence 31, Application PC/TUS0431036
; GENERAL INFORMATION:
; APPLICANT: Dale Umetsu
; APPLICANT: Rosemarie Dekruyff
; APPLICANT: Jennifer McIntire
; APPLICANT: Gordon Freeman
; TITLE OF INVENTION: T CELL REGULATORY GENES ASSOCIATED WITH
; TITLE OF INVENTION: IMMUNE DISEASE
; FILE REFERENCE: STAN-235CIP
; CURRENT APPLICATION NUMBER: PCT/US04/31036
; CURRENT FILING DATE: 2004-09-15
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 10/188,012
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 301
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(301)
; OTHER INFORMATION: TIM-3, allele 2
PCT-US04-31036-31

Query Match 12.3%; Score 71.5; DB 1; Length 301;
Best Local Similarity 28.0%; Pred. No. 3.4;
Matches 26; Conservative 12; Mismatches 28; Indels 27; Gaps 3;

Qy 10 LPGVAVVLLLLQSTQSVIYQGFVRVQLESMMKSLDLEAQWAPSPRILQAQSLPAVCHH 69
Db 5 LPFDCVLLLLLLLTSRSEVEYRA-----EVGNAYLPCF-YT 41
Qy 70 PALPDQLPVCASQAEASSIFK-----TLRTIAND 98
Db 42 PAAPGNLVPVCGKGCACPFCEGNNVLRTERD 74

RESULT 7
US-60-655-875-165482
; Sequence 165482, Application US/60655875
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETEROERA INFESTATIONS
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/60/655,875
; CURRENT FILING DATE: 2005-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 165482
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Heterodera glycines
```

```
; FEATURE:
; OTHER INFORMATION: Coding regions on vDNA: vDNA=SeqID_91905; Strand=+; Position=2-
US-60-655-875-165482

Query Match 11.8%; Score 68.5; DB 8; Length 166;
Best Local Similarity 24.0%; Pred. No. 3.6;
Matches 25; Conservative 19; Mismatches 41; Indels 19; Gaps 4;

Qy 21 LQSTQSVIYQGF-----FRVQLESMMKSLDLEAQWAPSPRILQAQSLPAVCHHPALPD 75
Db 48 LQKKSLHLSPRGTKSTHFSNEREAQKKQPKTKRRPQTTK--BRDALTKIHAQKLQ 105
Qy 76 LQPVCAQAEASS-----IFKTIIRTIANDDCELCVNVACTG 110
Db 106 LRRVQSDWESASVMVLRREGDAVVVKALR---QDGFSLCDNAVSSG 146

RESULT 8
US-10-450-763-46026
; Sequence 46026, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 46026
; LENGTH: 1444
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (996)..(1012)
; OTHER INFORMATION: Kw ANNEXIN ANTIGEN PROLINE TUMOR domain identified by
; OTHER INFORMATION: eMATRIX, accession number DM00250A, p-value=6.211e-10, raw score
; OTHER INFORMATION: 10.52
US-10-450-763-46026

Query Match 11.8%; Score 68.5; DB 6; Length 1444;
Best Local Similarity 31.4%; Pred. No. 47;
Matches 32; Conservative 9; Mismatches 46; Indels 15; Gaps 4;

Qy 4 RAASGLLPGVAVVLLLLQSTQSVIYQGFVRVQLESMMKSL-----DLAQWAPS 54
Db 1036 RARSGLOPQPQPAVLISLLGSAQ---VPOQG--VQLPSVLQQQPPKQLQFQWPK 1090
Qy 55 PRL-QAQSLLPAVCHHPALPDQLPVCASQAEASSIFKTLRTI 95
Db 1091 PRQEPQSQQQQPHIQLOTLQRLVLPQPHIQLOTLQQLRVL 1132

RESULT 9
US-11-031-175-15516
; Sequence 15516, Application US/11031175
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/11/031,175
; CURRENT FILING DATE: 2005-01-08
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
```



```
; APPLICANT: DASH, DEBASIS
; APPLICANT: SHARMA, RAMAKANT
; APPLICANT: MAHESHWARI, JITENDRA KUMAR
; TITLE OF INVENTION: A COMPUTER BASED VERSATILE METHOD FOR IDENTIFYING PROTEIN CODING
; TITLE OF INVENTION: DNA SEQUENCES USEFUL AS DRUG TARGETS
; FILE REFERENCE: 026033-00029
; CURRENT APPLICATION NUMBER: US/10/755,415
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: 10/727,989
; PRIOR FILING DATE: 2003-12-05
; NUMBER OF SEQ ID NOS: 373
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 236
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-755-415-236

Query Match      11.6%; Score 67; DB 6; Length 187;
Best Local Similarity 31.0%; Pred. No. 6;
Matches 18; Conservative 7; Mismatches 21; Indels 12; Gaps 2;

Qy 50 QWAPSPR---LQAQSLPA-----VCHHPALPQDLQPVCAEQEASSIFKTLRTI 95
Db 110 RWMSPRNRAIRALMALPAETALRSMQIGICTSPAMGSQVRPLCSMAISAAPSTCRGV 167

RESULT 14
US-11-033-545-409
; Sequence 409, Application US/11033545
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/11/033,545
; CURRENT FILING DATE: 2005-01-12
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 409
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Human
US-11-033-545-409

Query Match      11.5%; Score 66.5; DB 7; Length 228;
Best Local Similarity 32.0%; Pred. No. 8.7;
Matches 32; Conservative 12; Mismatches 35; Indels 21; Gaps 5;

Qy 5 AASGLPCVAVVLLVLLLOSTQSVYIQGFRVQLESMMKLLDLAQWAPSPRL----- 57
Db 104 AASGLPVLVVVLLVLL---AVLVTLAGIIVYRKARSRI--LSRNVAPKTTMGSRNPLF 158

Qy 58 -QAQSLPAPVCHHPAL---PDQLQVPC-----ASQEASSI 88
Db 159 HQAASRVPAKGAPAPSGPQELVPTTHPGQPARHPASSV 198

RESULT 15
US-10-489-448-1493
; Sequence 1493, Application US/10489448
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhou, Ping
```

```
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wang, Dunrui
; APPLICANT: Ma, Yunqing
; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Zhiwei
; APPLICANT: Weng, Gezhi
; APPLICANT: Haley-Vicente, Dana
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 810CIP PCT
; CURRENT APPLICATION NUMBER: US/10/489,448
; CURRENT FILING DATE: 1004-03-10
; PRIOR APPLICATION NUMBER: US 60/324,631
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: PCT/US00/35017
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: PCT/US01/03800
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 3476
; SOFTWARE: ptFL_Genes Version 6.0
; SEQ ID NO 1493
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-489-448-1493

Query Match      11.5%; Score 66.5; DB 6; Length 393;
Best Local Similarity 24.0%; Pred. No. 17;
Matches 24; Conservative 8; Mismatches 35; Indels 33; Gaps 2;

Qy 3 CRAASGLPGVAVVLLVLLLOSTQSVYIQGFRVQLESMMKLLSDLE----- 48
Db 242 CEVLKEVIGVVIETLDLVLET-----FRKSLRGQKMLPLLSQRRFVLLHNGEADP 292

Qy 49 -----AQWAPSPRLQAQSLPAPVCHHPALPQDLQ 78
Db 293 RPHLGSCSLRRPPLPTRQAKSGGPPMPHAPDQPGARRP 332

Search completed: March 26, 2005, 17:11:53
Job time : 34.7445 secs
```

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2005, 16:49:45 ; Search time 84.4821 Seconds  
(without alignments)  
86.982 Million cell updates/sec

Title: US-10-775-481A-2

Perfect score: 19  
Sequence: 1 NNTFYCELCCNPACAGCY 19

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003s:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 19    | 100.0       | 19     | 2     | AAR85946    |
| 2          | 19    | 100.0       | 19     | 2     | AAY40508    |
| 3          | 19    | 100.0       | 19     | 8     | ADR45822    |
| 4          | 19    | 100.0       | 72     | 8     | ADR48355    |
| 5          | 18    | 94.7        | 18     | 2     | AAR85956    |
| 6          | 18    | 94.7        | 18     | 2     | AAR85950    |
| 7          | 18    | 94.7        | 18     | 2     | AAY40543    |
| 8          | 18    | 94.7        | 18     | 2     | AAY40518    |
| 9          | 18    | 94.7        | 18     | 2     | AAY40512    |
| 10         | 18    | 94.7        | 18     | 2     | AAY02385    |
| 11         | 18    | 94.7        | 18     | 2     | AAY29607    |
| 12         | 18    | 94.7        | 18     | 2     | AAY02397    |
| 13         | 18    | 94.7        | 18     | 8     | ADR48330    |
| 14         | 18    | 94.7        | 18     | 8     | ADR48332    |
| 15         | 18    | 94.7        | 18     | 8     | ADR48341    |
| 16         | 18    | 94.7        | 18     | 8     | ADR45832    |
| 17         | 18    | 94.7        | 18     | 8     | ADR45857    |
| 18         | 18    | 94.7        | 18     | 8     | ADR45826    |
| 19         | 18    | 94.7        | 36     | 1     | AAP30262    |
| 20         | 17    | 89.5        | 17     | 2     | AAR85951    |
| 21         | 17    | 89.5        | 17     | 2     | AAR85957    |
| 22         | 17    | 89.5        | 17     | 2     | AAY40519    |
| 23         | 17    | 89.5        | 17     | 2     | AAY40546    |
| 24         | 17    | 89.5        | 17     | 2     | AAY40513    |
| 25         | 17    | 89.5        | 17     | 8     | ADR45827    |

|    |    |      |    |   |          |                    |
|----|----|------|----|---|----------|--------------------|
| 26 | 17 | 89.5 | 17 | 8 | ADR45833 | Adr45833 Amino aci |
| 27 | 17 | 89.5 | 17 | 8 | ADR45860 | Adr45860 Amino aci |
| 28 | 17 | 89.5 | 18 | 2 | AAY06971 | Aay06971 E. coli h |
| 29 | 16 | 84.2 | 16 | 2 | AAR85952 | Aar85952 ST Ia rec |
| 30 | 16 | 84.2 | 16 | 2 | AAR85958 | Aar85958 ST Ia rec |
| 31 | 16 | 84.2 | 16 | 2 | AAY40514 | Aay40514 ST recept |
| 32 | 16 | 84.2 | 16 | 2 | AAY40520 | Aay40520 ST recept |
| 33 | 16 | 84.2 | 16 | 8 | ADR45834 | Adr45834 Amino aci |
| 34 | 16 | 84.2 | 16 | 8 | ADR45828 | Adr45828 Amino aci |
| 35 | 15 | 78.9 | 15 | 2 | AAR85959 | Aar85959 ST Ia rec |
| 36 | 15 | 78.9 | 15 | 2 | AAR85953 | Aar85953 ST Ia rec |
| 37 | 15 | 78.9 | 15 | 2 | AAY40515 | Aay40515 ST recept |
| 38 | 15 | 78.9 | 15 | 2 | AAY40547 | Aay40547 ST recept |
| 39 | 15 | 78.9 | 15 | 2 | AAY40521 | Aay40521 ST recept |
| 40 | 15 | 78.9 | 15 | 8 | ADR45829 | Adr45829 Amino aci |
| 41 | 15 | 78.9 | 15 | 8 | ADR45835 | Adr45835 Amino aci |
| 42 | 15 | 78.9 | 15 | 8 | ADR45861 | Adr45861 Amino aci |
| 43 | 15 | 78.9 | 18 | 8 | ADR48342 | Adr48342 Bacterial |
| 44 | 14 | 73.7 | 14 | 2 | AAR85960 | Aar85960 ST Ia rec |
| 45 | 14 | 73.7 | 14 | 2 | AAR85954 | Aar85954 ST Ia rec |

#### ALIGNMENTS

RESULT 1  
AAR85946  
ID AAR85946 standard; peptide; 19 AA.  
XX  
AC AAR85946;  
XX  
DT 19-JAN-1996 (first entry)  
XX  
DE ST Ia receptor ligand portion.  
XX  
KW ST; receptor; ligand; heat-stable toxin; conjugate; imaging; cancer;  
KW colorectal; metastasis.  
XX  
OS Escherichia coli.  
XX  
PN W09511694-A1.  
XX  
PD 04-MAY-1995.  
XX  
PF 26-OCT-1994; 94WO-US012232.  
XX  
PR 26-OCT-1993; 93US-00141892.  
PR 13-SEP-1994; 94US-00305056.  
XX  
(UYJE-) UNIV JEFFERSON THOMAS.  
Waldman SA;  
XX  
DR WPI; 1995-178646/23.  
XX  
Conjugated cpds. which specifically bind to colorectal cancer cells -  
comprise heat-stable toxin receptor binding moiety and active moiety  
which may be a therapeutic agent or a radioactive agent.  
XX  
Claim 3; Page 116; 133pp; English.  
XX  
New conjugated compounds are provided which consist of (1) an ST receptor  
binding moiety and (2) an active moiety which is a radio- stable agent.  
XX  
'ST' refers to a heat stable toxin produced by E.coli and other  
XX  
organisms. Especially the radiostable agent is a therapeutic agent (e.g.  
XX  
methotrexate, doxorubicin, mitomycin), imaging agent (e.g. radioactive  
XX  
isotope) or nucleic acid; and the compound is used for the detection,  
XX  
imaging or treatment of colorectal tumours, particularly metastasised  
XX  
tumours. The present sequence is a specific example of an ST receptor  
XX  
binding peptide which can be used in the conjugate  
SQ Sequence 19 AA;

```
Query Match      100.0%; Score 19; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 1e-13;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNTFYCCELCNPPACAGCY 19
DB 1 NNTFYCCELCNPPACAGCY 19

RESULT 2
AAY40508
ID AAY40508 standard; peptide; 19 AA.
XX
AC AAY40508;
XX
DT 03-DEC-1999 (first entry)
XX
DE ST receptor peptide ST 1a.
XX
KW Conjugated compound; ST receptor; heat-stable; toxin; colorectal cancer.
XX
OS Escherichia coli.
XX
XX US5962220-A.
XX
PD 05-OCT-1999.
XX
XX 06-JUN-1995; 95US-00467920.
XX
XX 26-OCT-1993; 93US-00141892.
XX
XX (UYJE-) UNIV JEFFERSON THOMAS.
XX
XX Waldman SA;
XX
XX WPI; 1999-571264/48.
XX
DR N-PSDB; AAZ07540.
XX
XX Conjugated compound comprising a receptor moiety and active moiety,
XX useful for the treatment/prevention of colorectal cancer.
XX
XX Claim 1; Col 19-20; 23pp; English.
XX
XX The invention relates to a conjugated compound that comprises a ST (heat-
XX stable toxin) receptor moiety selected from one of the sequences shown in
XX AAY40508-Y40559 and an active moiety (antisense molecule). The compound
XX is useful for the treatment of colorectal cancer. The use of the
XX conjugated compound is advantageous compared to antibodies since it binds
XX specifically to colorectal cells and has no toxic effect on normal cells.
XX The present sequence represents the amino acid sequence of a ST receptor
XX peptide ST 1a
XX
SQ Sequence 19 AA;
Query Match      100.0%; Score 19; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 1e-13;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNTFYCCELCNPPACAGCY 19
DB 1 NNTFYCCELCNPPACAGCY 19

RESULT 3
ADR45822
ID ADR45822 standard; peptide; 19 AA.
XX
AC ADR45822;
XX
DT 18-NOV-2004 (first entry)
XX
XX Amino acid sequence of heat stable toxin ST 1a.
XX
```

```
KW heat stable toxin; ST; ST receptor; colorectal cancer; gastric cancer;
KW oesophageal cancer; colorectal cell; gastric cell; oesophageal cell;
XX ST 1a.
XX
OS Unidentified.
XX
XX WO2004071436-A2.
XX
PD 26-AUG-2004.
XX
XX 10-FEB-2004; 2004WO-US003765.
XX
XX 10-FEB-2003; 2003US-0446730P.
XX
XX (UYJE-) UNIV JEFFERSON THOMAS.
XX
XX Waldman SA, Pitari GM, Park J, Schulz S, Wolfe HR;
XX
XX WPI; 2004-615913/59.
XX
DR N-PSDB; ADR45821.
XX
XX Increasing ST receptor molecules on the surface of a colorectal, gastric
XX or esophageal cancer cell to treat these cancers comprises administering
XX ST receptor ligand molecules that bind to ST receptors on the surface of
XX the cancer cell.
XX
XX Claim 6; SEQ ID NO 2; 97pp; English.
XX
XX The specification describes a method for increasing the number of heat
XX stable toxin (ST) receptor molecules on the surface of a metastasised
XX colorectal cancer cell. The method comprises administering, by continuous
XX infusion, at least 0.1 nM of an ST receptor ligand per 10 kg bodyweight
XX per hour for at least 6 hours, where ST receptor ligand molecules bind to
XX ST receptors on the surface of the cancer cell in the individual and the
XX number of ST receptor molecules on the surface of the cancer cell is
XX increased. Therapeutic compositions comprising components which target ST
XX receptors can then be used to inhibit proliferation of the colorectal,
XX gastric and oesophageal cells. This method may be used for treating
XX individuals that have diseases that affect colorectal, gastric and
XX oesophageal cells, including colorectal, gastric or oesophageal cancers.
XX The present sequence represents a ST, designated ST 1a, which is used as
XX the ST receptor ligand in the method of the invention.
XX
SQ Sequence 19 AA;
Query Match      100.0%; Score 19; DB 8; Length 19;
Best Local Similarity 100.0%; Pred. No. 1e-13;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNTFYCCELCNPPACAGCY 19
DB 1 NNTFYCCELCNPPACAGCY 19

RESULT 4
ADR48355
ID ADR48355 standard; peptide; 72 AA.
XX
XX ADR48355;
XX
XX 04-NOV-2004 (first entry)
XX
XX Immature E. coli ST-1A peptide.
XX
XX Gastrointestinal; antiinflammatory; laxative; cardiant; antilucer;
XX anorectic; cardiovascular; cytostatic; analgesic; CNS; respiratory;
XX neuroprotective; vasotropic; auditory; antiemetic; antiasthmatic;
XX nephrotropic; hepatotropic; virucide; immunosuppressive; antiallergic;
XX antidiabetic; ophthalmological; tranquiliser; hypnotic; nootropic;
XX guanilate cyclase C; GC-C; receptor; gastrointestinal disorder;
XX irritable bowel syndrome; constipation; gastroesophageal reflux disease;
XX heartburn; dyspepsia; gastroparesis; Crohn's disease; ulcerative colitis;
XX inflammatory bowel disease; obesity; heart failure; cystic fibrosis;
```

KW cancer; respiratory disorder; neurological disorder; carbonate imbalance;  
KW erectile dysfunction; inner ear disorder; slow digestion; nausea;  
KW vomiting; bloating; asthma; nephritis; hepatitis; pancreatitis; allergy;  
KW retinopathy; nephropathy; headache; anxiety; sleep disorder.

XX Escherichia coli.

XX WO2004069165-A2.

XX 19-AUG-2004.

XX 28-JAN-2004; 2004WO-US002390.

XX 28-JAN-2003; 2003US-0443098P.

PR 15-MAY-2003; 2003US-0471288P.

XX 12-NOV-2003; 2003US-0519460P.

XX (MICR-) MICROBIA INC.

XX Currie MG, Mahajan-Miklos S;

PI WPI; 2004-604332/58.

XX Novel purified peptide capable of activating the guanylate cyclase C  
XX receptor, useful for treating obesity, congestive heart failure and  
XX benign prostatic hyperplasia.

PS Disclosure; Page 30; 93pp; English.

XX The invention relates to a purified peptide (P1) capable of activating  
CC the guanylate cyclase C (GC-C) receptor. Further disclosed is a  
CC pharmaceutical composition comprising the peptide of the invention. The  
CC composition of the invention is useful for treating a gastrointestinal  
CC disorder in a patient, which involves administering P1, where the  
CC gastrointestinal disorder is gastrointestinal motility disorder,  
CC irritable bowel syndrome, chronic constipation, a functional  
CC gastrointestinal disorder, gastroesophageal reflux disease, functional  
CC heartburn, dyspepsia, functional dyspepsia, nonulcer dyspepsia,  
CC gastroparesis, chronic intestinal pseudo-obstruction, colonic pseudo-  
CC obstruction, Crohn's disease, ulcerative colitis or inflammatory bowel  
CC disease. The peptide of the invention is also useful for treating  
CC obesity, congestive heart failure, cystic fibrosis or a patient suffering  
CC from constipation. The P1/GC-C receptor agonist is useful for treating  
CC cancer, respiratory disorder, neurological disorder, disorder associated  
CC with carbonate imbalance, erectile dysfunction, insulin-related disorder  
CC or inner ear disorder. P1 is useful in treating slow digestion or slow  
CC stomach emptying. P1 is useful in relieving symptoms of gastroparesis,  
CC such as nausea, vomiting, bloating, and delayed gastric emptying. P1 is  
CC useful for treating or preventing asthma, nephritis, hepatitis,  
CC pancreatitis, allergies, etc. P1 is useful for treating or preventing  
CC including inhalation. P1 is useful in treating or preventing  
CC nephropathy and edema formation. P1 is useful for treating or preventing  
CC headache, anxiety, sleep disorders and memory loss. P1 is useful as a  
CC marker to identify, detect, stage, or diagnosis diseases and conditions  
CC of the small intestine, including Crohn's disease, colitis, inflammatory  
CC bowel disease, tumors, etc. P1 can be conjugated to diagnostic or  
CC therapeutic molecule to target cells bearing GC-C receptor, e.g., cystic  
CC fibrosis lesions and specific cells lining the intestinal tract, thus  
CC useful in targeting radioactive moieties or therapeutic moieties to the  
CC intestine to aid in imaging and diagnosing or treating  
CC colorectal/metastatic or local colorectal cancer. The current sequence  
CC represents a bacterial ST peptide which is an activator of the GC-C  
CC receptor. ST peptides are considered super agonists of GC-C and are very  
CC resistant to proteolytic degradation.

XX Sequence 72 AA;

Query Match 100.0%; Score 19; DB 8; Length 72;  
Best Local Similarity 100.0%; Pred. No. 2.9e-13;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NNTFYCCCLCCNPACAGCY 19

Db 54 NNTFYCCCLCCNPACAGCY 72

RESULT 5  
AAR85956

ID AAR85956 standard; peptide; 18 AA.

XX AAR85956;

XX 19-JAN-1996 (first entry)

XX ST Ia receptor ligand portion.

XX ST; receptor; ligand; heat-stable toxin; conjugate; imaging; cancer;  
XX colorectal; metastasis.

XX Escherichia coli.

XX WO9511694-A1.

XX 04-MAY-1995.

XX 26-OCT-1994; 94WO-US012232.

XX 26-OCT-1993; 93US-00141892.

XX 13-SEP-1994; 94US-00305056.

XX (UWJE-) UNIV JEFFERSON THOMAS.

XX Waldman SA;

XX WPI; 1995-178646/23.

XX Conjugated cpds. which specifically bind to colorectal cancer cells -  
XX comprise heat-stable toxin receptor binding moiety and active moiety  
XX which may be a therapeutic agent or a radioactive agent.

XX Claim 3; Page 118; 133pp; English.

XX New conjugated compounds are provided which consist of (1) an ST receptor  
XX binding moiety and (2) an active moiety which is a radio-stable agent.  
XX 'ST' refers to a heat stable toxin produced by E.coli and other  
XX organisms. Especially the radiostable agent is a therapeutic agent (e.g.  
XX methotrexate, doxorubicin, mitomycin), imaging agent (e.g. radioactive  
XX isotope) or nucleic acid; and the compound is used for the detection,  
XX imaging or treatment of colorectal tumours, particularly metastasised  
XX tumours. The present sequence is a specific example of an ST receptor  
XX binding peptide which can be used in the conjugate

XX Sequence 18 AA;

Query Match 94.7%; Score 18; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 1.2e-12;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 NTFYCCCLCCNPACAGCY 19

Db 1 NTFYCCCLCCNPACAGCY 18

RESULT 6  
AAR85950

ID AAR85950 standard; peptide; 18 AA.

XX AAR85950;

XX 19-JAN-1996 (first entry)

XX ST Ia receptor ligand portion.

XX ST; receptor; ligand; heat-stable toxin; conjugate; imaging; cancer;  
XX colorectal; metastasis.

```

XX OS Escherichia coli.
XX PN WO9511694-A1.
XX PD 04-MAY-1995.
XX XX 26-OCT-1994; 94WO-US012232.
XX PF 26-OCT-1993; 93US-00141892.
XX PR 13-SEP-1994; 94US-00305056.
XX PA (UYJE-) UNIV JEFFERSON THOMAS.
XX PI Waldman SA;
XX XX WPI; 1995-178646/23.
XX DR Conjugated cpds, which specifically bind to colorectal cancer cells -
XX PT comprise heat-stable toxin receptor binding moiety and active moiety
XX PT which may be a therapeutic agent or a radioactive agent.
XX PS Claim 3; Page 117; 133pp; English.
XX CC New conjugated compounds are provided which consist of (1) an ST receptor
XX CC binding moiety and (2) an active moiety which is a radio- stable agent.
XX CC 'ST' refers to a heat stable toxin produced by E.coli and other
XX CC organisms. Especially the radiostable agent is a therapeutic agent (e.g.
XX CC methotrexate, doxorubicin, mitomycin), imaging agent (e.g. radioactive
XX CC isotope) or nucleic acid; and the compound is used for the detection,
XX CC imaging or treatment of colorectal tumors, particularly metastasised
XX CC tumours. The present sequence is a specific example of an ST receptor
XX CC binding peptide which can be used in the conjugate
XX SQ Sequence 18 AA;

Query Match 94.7%; Score 18; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.2e-12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNTFYCCELCNCPACAGC 18
Db 1 NNTFYCCELCNCPACAGC 18

RESULT 7
AAAY40543
ID AAY40543 standard; peptide; 18 AA.
XX AC AAY40543;
XX DT 03-DEC-1999 (first entry)
XX DE ST receptor binding peptide.
XX KW Conjugated compound; ST receptor; heat-stable; toxin; colorectal cancer.
XX OS Escherichia coli.
XX PN US5962220-A.
XX PD 05-OCT-1999.
XX PF 06-JUN-1995; 95US-00467920.
XX PR 26-OCT-1993; 93US-00141892.
XX PA (UYJE-) UNIV JEFFERSON THOMAS.
XX PI Waldman SA;
XX XX WPI; 1999-571264/48.
XX DR Conjugated compound comprising a receptor moiety and active moiety,
XX PT useful for the treatment/prevention of colorectal cancer.
XX PS Claim 1; Col 25-26; 23pp; English.
XX CC The invention relates to a conjugated compound that comprises a ST (heat-
XX CC stable toxin) receptor moiety selected from one of the sequences shown in
XX CC AAY40508-Y40559 and an active moiety (antitense molecule). The compound
XX CC is useful for the treatment of colorectal cancer. The use of the
XX CC conjugated compound is advantageous compared to antibodies since it binds
XX CC specifically to colorectal cells and has no toxic effect on normal cells
XX SQ Sequence 18 AA;

Query Match 94.7%; Score 18; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.2e-12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNTFYCCELCNCPACAGCY 19
Db 1 NNTFYCCELCNCPACAGCY 18

RESULT 8
AAAY40518
ID AAY40518 standard; peptide; 18 AA.
XX AC AAY40518;
XX DT 03-DEC-1999 (first entry)
XX DE ST receptor binding peptide.
XX KW Conjugated compound; ST receptor; heat-stable; toxin; colorectal cancer.
XX OS Escherichia coli.
XX PN US5962220-A.
XX PD 05-OCT-1999.
XX PF 06-JUN-1995; 95US-00467920.
XX PR 26-OCT-1993; 93US-00141892.
XX PA (UYJE-) UNIV JEFFERSON THOMAS.
XX PI Waldman SA;
XX XX WPI; 1999-571264/48.
XX DR Conjugated compound comprising a receptor moiety and active moiety,
XX PT useful for the treatment/prevention of colorectal cancer.
XX PS Claim 1; Col 25-26; 23pp; English.
XX CC The invention relates to a conjugated compound that comprises a ST (heat-
XX CC stable toxin) receptor moiety selected from one of the sequences shown in
XX CC AAY40508-Y40559 and an active moiety (antitense molecule). The compound
XX CC is useful for the treatment of colorectal cancer. The use of the
XX CC conjugated compound is advantageous compared to antibodies since it binds
XX CC specifically to colorectal cells and has no toxic effect on normal cells
XX SQ Sequence 18 AA;

Query Match 94.7%; Score 18; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.2e-12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNTFYCCELCNCPACAGCY 19
Db 1 NNTFYCCELCNCPACAGCY 18

```



RESULT 9  
 AAY40512  
 ID AAY40512 standard; peptide; 18 AA.  
 XX  
 AC AAY40512;  
 DT 03-DEC-1999 (first entry)  
 XX ST receptor binding peptide.  
 DE  
 XX Conjugated compound; ST receptor; heat-stable; toxin; colorectal cancer.  
 KW Escherichia coli.  
 XX OS  
 XX US9562220-A.  
 XX PN  
 XX PD 05-OCT-1999.  
 XX PF 06-JUN-1995; 95US-00467920.  
 XX PR 26-OCT-1993; 93US-00141892.  
 XX PA (UYJE-) UNIV JEFFERSON THOMAS.  
 XX PI Waldman SA;  
 XX DR WPI; 1999-571264/48.  
 XX PT Conjugated compound comprising a receptor moiety and active moiety,  
 PT useful for the treatment/prevention of colorectal cancer.  
 XX PS Claim 1; Col 21-22; 23pp; English.  
 XX CC The invention relates to a conjugated compound that comprises a ST (heat-  
 CC stable toxin) receptor moiety selected from one of the sequences shown in  
 CC AAY40508-Y40559 and an active moiety (antitense molecule). The compound  
 CC is useful for the treatment of colorectal cancer. The use of the  
 CC conjugated compound is advantageous compared to antibodies since it binds  
 CC specifically to colorectal cells and has no toxic effect on normal cells  
 XX  
 SQ Sequence 18 AA;  
 Query Match 94.7%; Score 18; DB 2; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-12;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 NNTFYCCELCNCPACAGC 18  
 Db 1 NNTFYCCELCNCPACAGC 18  
 RESULT 10  
 AAY02385  
 ID AAY02385 standard; peptide; 18 AA.  
 XX  
 AC AAY02385;  
 DT 09-JUL-1999 (first entry)  
 XX Heat stable ST enterotoxin Sta.  
 DE  
 XX Selection; candidate drug; cell receptor binding; affinity;  
 KW biological receptor; rational drug design; combinatorial drug design;  
 KW receptor antagonist; receptor agonist; ST enterotoxin; beta turn mimetic;  
 KW gamma-turn mimetic; beta sheet mimetic; disulphide bridge mimetic.  
 XX OS  
 XX Escherichia coli.  
 XX PN WO9909416-A2.  
 XX PD 25-FEB-1999.  
 XX PS  
 PF 20-AUG-1998; 98WO-GB002504.  
 XX PR 20-AUG-1997; 97GB-00017652.  
 XX PA (NYCO-) NYCOMED IMAGING AS.  
 XX PA (COCK/) COCKBAIN J.  
 XX PI Wolfe HR;  
 XX DR WPI; 1999-181156/15.  
 XX PT Method of drug selection - and use of an acetamidomethyl-protected  
 PT polymer as a substrate in the solid state synthesis of an oligopeptide.  
 XX PS Disclosure; Page 1; 38pp; English.  
 XX CC The specification describes a method for selecting a candidate drug  
 CC compound having affinity for biological receptors. The method uses a  
 CC combination of rational and combinatorial drug design techniques. At  
 CC least 1 residue in the original cell receptor binding peptide is modified  
 CC to a non-natural amino acid, preferably a beta turn mimetic, a gamma-turn  
 CC mimetic, a beta sheet mimetic or a disulphide bridge mimetic. The method  
 CC is used for identification of a candidate receptor antagonist or agonist.  
 CC The present peptide is a cell receptor binding peptide, and can thus be  
 CC used as a starting point for identification of candidate drug compounds,  
 CC using the method of the invention  
 XX  
 SQ Sequence 18 AA;  
 Query Match 94.7%; Score 18; DB 2; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-12;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 2 NTFYCCCELCNCPACAGCY 19  
 Db 1 NTFYCCCELCNCPACAGCY 18  
 RESULT 11  
 AAY29607  
 ID AAY29607 standard; peptide; 18 AA.  
 XX  
 AC AAY29607;  
 DT 15-OCT-1999 (first entry)  
 XX Escherichia coli heat stable ST enterotoxin Sta.  
 DE Heat stable ST enterotoxin; immunoreagent; radiological therapy;  
 KW diagnosis; ST receptor binding moiety; macrocyclic complexing agent;  
 KW tumour; infectious diarrhoeal disease; diarrhoea.  
 XX OS  
 XX Escherichia coli.  
 XX PN WO9939748-A1.  
 XX PD 12-AUG-1999.  
 XX PF 08-FEB-1999; 99WO-GB000396.  
 XX PR 06-FEB-1998; 98US-00020233.  
 XX PA (NYCO-) NYCOMED IMAGING AS.  
 XX PA (MATY/) MATTHEWS D P.  
 XX PI Snow RA, Delecki DJ, Shah C, Black C, Wolfe H;  
 XX DR WPI; 1999-494219/41.  
 XX PT Macrocyclic complexing agents containing linked 2,6-pyridinylene nuclei  
 PT as components of targeting immunoreagents binding to ST receptor.  
 XX PS Disclosure; Page 39; 79pp; English.

XX The present invention describes targeting immunoreagents (TI's)  
CC comprising a metal ion and a residue of a macrocyclic complexing agent  
CC (MCA). TI's are of use in diagnostic imaging and therapy of specific  
CC disease sites in a patient, using either radioactive, magnetic resonance,  
CC or fluorescent means of detection or use of the metal ion; alternatively,  
CC a substituent of these types may be introduced, e.g. radioactive iodine,  
CC to perform the same function. Most notable is the imaging and  
CC radiological therapy of tumours. In addition, a variety of bacteria,  
CC including *Escherichia coli*, *Vibrio cholerae*, *Citrobacter freundii*, and  
CC *Yersinia enterocolitica*, bind to ST receptors and cause infectious  
CC diarrhoeal diseases, particularly in pediatrics and in developing  
CC countries. These types of diarrhoea can also be treated using TI's. TI's  
CC may specifically used to treat cancers and also be used as an  
CC anti-diarrhoeal agent. TI's are free from the various disadvantages of  
CC prior art reagents, including rapid destruction and/or excretion,  
CC instability in storage, and protein degradation. There is no perturbation  
CC of protein reactive groups at the pyridyl chelating site. AAY29607 to  
CC AAY29612 represent examples of heat stable ST enterotoxins given in the  
CC exemplification of the present invention  
XX  
XX Sequence 18 AA;

Query Match 94.7%; Score 18; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 1.2e-12;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NTFYCCCLCCNPACAGCY 19  
DB 1 NTFYCCCLCCNPACAGCY 18  
|||||

## RESULT 12

RAY02397  
ID AAY02397 standard; peptide; 18 AA.

XX AAY02397;

DT 09-JUL-1999 (first entry)

DE Heat stable ST enterotoxin Sta.

XX Selection; candidate drug; cell receptor binding; affinity;  
KW biological receptor; rational drug design; combinatorial drug design;  
KW receptor antagonist; receptor agonist; ST enterotoxin; beta turn mimetic;  
KW gamma-turn mimetic; beta sheet mimetic; disulphide bridge mimetic.

XX *Escherichia coli*.

XX WO9909417-A2.

XX 25-FEB-1999.

XX 20-AUG-1998; 98WO-GB002510.

XX 20-AUG-1997; 97GB-00017652.

XX (NYCO-) NYCOMED IMAGING AS.

PA (COCK/) COCKBAIN J.

XX Wolfe HR;

XX WPI; 1999-181157/15.

XX Method of drug selection - using a combination of rational and  
PT combinatorial drug design techniques.

XX Disclosure; Page 1; 35pp; English.

XX The specification describes a method for selecting a candidate drug  
CC compound having affinity for biological receptors. The method uses a  
CC combination of rational and combinatorial drug design techniques. At  
CC least 1 residue in the original cell receptor binding peptide is modified

CC to a non-natural amino acid, preferably a beta turn mimetic, a gamma-turn  
CC mimetic, a beta sheet mimetic or a disulphide bridge mimetic. The method  
CC is used for identification of a candidate receptor antagonist or agonist.  
CC The present peptide is a cell receptor binding peptide, and can thus be  
CC used as a starting point for identification of candidate drug compounds,  
XX using the method of the invention

XX Sequence 18 AA;

Query Match 94.7%; Score 18; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 1.2e-12;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NTFYCCCLCCNPACAGCY 19  
DB 1 NTFYCCCLCCNPACAGCY 18  
|||||

## RESULT 13

ADR48330

ID ADR48330 standard; peptide; 18 AA.

XX ADR48330;

DT 04-NOV-2004 (first entry)

DE *E. coli* ST Ia peptide.

XX Gastrointestinal; antiinflammatory; laxative; cardiant; antiulcer;  
KW anorectic; cardiovascular; cytostatic; analgesic; CNS; respiratory;  
KW neuroprotective; vasotropic; auditory; antiemetic; antiasthmatic;  
KW nephrotropic; hepatotropic; virucide; immunosuppressive; antiallergic;  
KW antidiabetic; ophthalmological; tranquiliser; hypnotic; nootropic;  
KW guanylate cyclase C; GC-C; receptor; gastrointestinal disorder;  
KW irritable bowel syndrome; constipation; gastroesophageal reflux disease;  
KW heartburn; dyspepsia; gastroparesis; Crohn's disease; ulcerative colitis;  
KW inflammatory bowel disease; obesity; heart failure; cystic fibrosis;  
KW cancer; respiratory disorder; neurological disorder; carbonate imbalance;  
KW erectile dysfunction; inner ear disorder; slow digestion; nausea;  
KW vomiting; bloating; asthma; nephritis; hepatitis; pancreatitis; allergy;  
KW retinopathy; nephropathy; headache; anxiety; sleep disorder.

XX *Escherichia coli*.

OS WO2004069165-A2.

XX 19-AUG-2004.

XX 28-JAN-2004; 2004WO-US002390.

XX 28-JAN-2003; 2003US-0443098P.

PR 15-MAY-2003; 2003US-0471288P.

PR 12-NOV-2003; 2003US-0519460P.

XX (MICR-) MICROBIA INC.

XX Currie MG, Mahajan-Miklos S;

XX WPI; 2004-604332/58.

XX Novel purified peptide capable of activating the guanylate cyclase C

PT receptor, useful for treating obesity, congestive heart failure and

PT benign prostatic hyperplasia.

XX Disclosure; Page 28; 93pp; English.

XX The invention relates to a purified peptide (PI) capable of activating  
CC the guanylate cyclase C (GC-C) receptor. Further disclosed is a  
CC pharmaceutical composition comprising the peptide of the invention. The  
CC composition of the invention is useful for treating a gastrointestinal  
CC disorder in a patient, which involves administering PI, where the  
CC gastrointestinal disorder is gastrointestinal motility disorder,  
CC irritable bowel syndrome, chronic constipation, a functional

CC gastrointestinal disorder, gastroesophageal reflux disease, functional  
 CC heartburn, dyspepsia, functional dyspepsia, nonulcer dyspepsia,  
 CC gastroparesis, chronic intestinal pseudo-obstruction, colonic pseudo-  
 CC obstruction, Crohn's disease, ulcerative colitis or inflammatory bowel  
 CC disease. The peptide of the invention is also useful for treating  
 CC obesity, congestive heart failure, cystic fibrosis or a patient suffering  
 CC from constipation. The P1/GC-C receptor agonist is useful for treating  
 CC cancer, respiratory disorder, neurological disorder, disorder associated  
 CC with carbonate imbalance, erectile dysfunction, insulin-related disorder  
 CC or inner ear disorder. P1 is useful in treating slow digestion or slow  
 CC stomach emptying. P1 is useful in relieving symptoms of gastroparesis  
 CC such as nausea, vomiting, bloating, and delayed gastric emptying. P1 is  
 CC useful for treating or preventing asthma, nephritis, hepatitis,  
 CC pancreatitis, allergies, etc. P1 is useful for treating or preventing  
 CC type II diabetes mellitus, hyperglycemia, respiratory disorders  
 CC including inhalation. P1 is useful in treating or preventing retinopathy,  
 CC nephropathy and edema formation. P1 is useful for treating or preventing  
 CC headache, anxiety, sleep disorders and memory loss. P1 is useful as a  
 CC marker to identify, detect, stage, or diagnosis diseases and conditions  
 CC of the small intestine, including Crohn's disease, colitis, inflammatory  
 CC bowel disease, tumours, etc. P1 can be conjugated to diagnostic or  
 CC therapeutic molecule to target cells bearing GC-C receptor, e.g., cystic  
 CC fibrosis lesions and specific cells lining the intestinal tract, thus  
 CC useful in targeting radioactive moieties or therapeutic moieties to the  
 CC intestine to aid in imaging and diagnosing or treating  
 CC colorectal/metastasised or local colorectal cancer. The current sequence  
 CC represents a bacterial ST peptide which is an activator of the GC-C  
 CC receptor. ST peptides are considered super agonists of GC-C and are very  
 CC resistant to proteolytic degradation.

XX Sequence 18 AA;

Query Match 94.7%; Score 18; DB 8; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-12;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NTFYCCCLCCNPACAGCY 19

Db 1 NTFYCCCLCCNPACAGCY 18

RESULT 14

ADR48332 94.7%; Score 18; DB 8; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-12;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX ADR48332;

XX 04-NOV-2004 (first entry)

DE C. freundii ST peptide.

KW Gastrointestinal; antiinflammatory; laxative; cardiant; antiulcer;  
 KW anorectic; cardiovascular; cytostatic; analgesic; CNS; respiratory;  
 KW neuroprotective; vasotropic; auditory; antiemetic; antidiabetic;  
 KW nephrotropic; hepatotropic; virucide; immunosuppressive; antiallergic;  
 KW antidiabetic; ophthalmological; tranquiliser; hypnotic; nootropic;  
 KW guanylate cyclase C; GC-C; receptor; gastrointestinal disorder;  
 KW irritable bowel syndrome; constipation; gastroesophageal reflux disease;  
 KW heartburn; dyspepsia; gastroparesis; Crohn's disease; ulcerative colitis;  
 KW inflammatory bowel disease; obesity; heart failure; cystic fibrosis;  
 KW cancer; respiratory disorder; neurological disorder; carbonate imbalance;  
 KW erectile dysfunction; inner ear disorder; slow digestion; nausea;  
 KW vomiting; bloating; asthma; nephritis; hepatitis; pancreatitis; allergy;  
 KW retinopathy; nephropathy; headache; anxiety; sleep disorder.

XX Citrobacter freundii.

XX WO2004069165-A2.

XX 19-AUG-2004.

XX 28-JAN-2004; 2004WO-US002390.

XX

PR 28-JAN-2003; 2003US-0443098P.

PR 15-MAY-2003; 2003US-0471288P.

XX 12-NOV-2003; 2003US-0519460P.

PA (MICR-) MICROBIA INC.

XX Currie MG, Mahajan-Miklos S;

XX WPI; 2004-604332/58.

XX Novel purified peptide capable of activating the guanylate cyclase C  
 XX receptor, useful for treating obesity, congestive heart failure and  
 XX benign prostatic hyperplasia.

XX Disclosure; Page 28; 93pp; English.

XX The invention relates to a purified peptide (P1) capable of activating  
 XX the guanylate cyclase C (GC-C) receptor. Further disclosed is a  
 XX pharmaceutical composition comprising the peptide of the invention. The  
 XX composition of the invention is useful for treating a gastrointestinal  
 XX disorder in a patient, which involves administering P1, where the  
 XX gastrointestinal disorder is gastrointestinal motility disorder,  
 XX irritable bowel syndrome, chronic constipation, a functional  
 XX gastrointestinal disorder, gastroesophageal reflux disease, functional  
 XX heartburn, dyspepsia, functional dyspepsia, nonulcer dyspepsia,  
 XX gastroparesis, chronic intestinal pseudo-obstruction, colonic pseudo-  
 XX obstruction, Crohn's disease, ulcerative colitis or inflammatory bowel  
 XX disease. The peptide of the invention is also useful for treating  
 XX obesity, congestive heart failure, cystic fibrosis or a patient suffering  
 XX from constipation. The P1/GC-C receptor agonist is useful for treating  
 XX cancer, respiratory disorder, neurological disorder, disorder associated  
 XX with carbonate imbalance, erectile dysfunction, insulin-related disorder  
 XX or inner ear disorder. P1 is useful in treating slow digestion or slow  
 XX stomach emptying. P1 is useful in relieving symptoms of gastroparesis  
 XX such as nausea, vomiting, bloating, and delayed gastric emptying. P1 is  
 XX useful for treating or preventing asthma, nephritis, hepatitis,  
 XX pancreatitis, allergies, etc. P1 is useful for treating or preventing  
 XX type II diabetes mellitus, hyperglycemia, respiratory disorders  
 XX including inhalation. P1 is useful in treating or preventing retinopathy,  
 XX nephropathy and edema formation. P1 is useful for treating or preventing  
 XX headache, anxiety, sleep disorders and memory loss. P1 is useful as a  
 XX marker to identify, detect, stage, or diagnosis diseases and conditions  
 XX of the small intestine, including Crohn's disease, colitis, inflammatory  
 XX bowel disease, tumours, etc. P1 can be conjugated to diagnostic or  
 XX therapeutic molecule to target cells bearing GC-C receptor, e.g., cystic  
 XX fibrosis lesions and specific cells lining the intestinal tract, thus  
 XX useful in targeting radioactive moieties or therapeutic moieties to the  
 XX intestine to aid in imaging and diagnosing or treating  
 XX colorectal/metastasised or local colorectal cancer. The current sequence  
 XX represents a bacterial ST peptide which is an activator of the GC-C  
 XX receptor. ST peptides are considered super agonists of GC-C and are very  
 XX resistant to proteolytic degradation.

XX Sequence 18 AA;

Query Match 94.7%; Score 18; DB 8; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-12;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NTFYCCCLCCNPACAGCY 19

Db 1 NTFYCCCLCCNPACAGCY 18

RESULT 15

ADR48341

ID ADR48341 standard; peptide; 18 AA.

XX ADR48341;

XX 04-NOV-2004 (first entry)

XX Bacterial ST peptide #2.

DE



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2005, 17:12:01 ; Search time 25.7857 Seconds  
(without alignments)  
55.005 Million cell updates/sec

Title: US-10-775-481A-2

Perfect score: 19

Sequence: 1 NNTFYCCELCCNPACAGCY 19

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents\_AA.\*

- 1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pcp.\*
- 2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pcp.\*
- 3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pcp.\*
- 4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pcp.\*
- 5: /cgn2\_6/ptodata/1/1aa/PTCUS\_COMB.pcp.\*
- 6: /cgn2\_6/ptodata/1/1aa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 19    | 100.0       | 19     | 1  | US-08-141-892A-2  |
| 2          | 19    | 100.0       | 19     | 2  | US-08-583-447A-2  |
| 3          | 19    | 100.0       | 19     | 2  | US-08-467-920-2   |
| 4          | 19    | 100.0       | 19     | 3  | US-08-635-930-2   |
| 5          | 19    | 100.0       | 19     | 3  | US-09-193-997-2   |
| 6          | 19    | 100.0       | 19     | 3  | US-09-138-237A-2  |
| 7          | 18    | 94.7        | 18     | 1  | US-08-141-892A-7  |
| 8          | 18    | 94.7        | 18     | 1  | US-08-141-892A-13 |
| 9          | 18    | 94.7        | 18     | 1  | US-08-141-892A-38 |
| 10         | 18    | 94.7        | 18     | 2  | US-08-583-447A-7  |
| 11         | 18    | 94.7        | 18     | 2  | US-08-583-447A-13 |
| 12         | 18    | 94.7        | 18     | 2  | US-08-583-447A-38 |
| 13         | 18    | 94.7        | 18     | 2  | US-08-467-920-7   |
| 14         | 18    | 94.7        | 18     | 2  | US-08-467-920-13  |
| 15         | 18    | 94.7        | 18     | 2  | US-08-467-920-38  |
| 16         | 18    | 94.7        | 18     | 3  | US-08-635-930-7   |
| 17         | 18    | 94.7        | 18     | 3  | US-08-635-930-13  |
| 18         | 18    | 94.7        | 18     | 3  | US-08-635-930-38  |
| 19         | 18    | 94.7        | 18     | 3  | US-09-193-997-7   |
| 20         | 18    | 94.7        | 18     | 3  | US-09-193-997-13  |
| 21         | 18    | 94.7        | 18     | 3  | US-09-193-997-38  |
| 22         | 18    | 94.7        | 18     | 3  | US-09-138-237A-7  |
| 23         | 18    | 94.7        | 18     | 3  | US-09-138-237A-13 |
| 24         | 18    | 94.7        | 18     | 3  | US-09-138-237A-38 |
| 25         | 17    | 89.5        | 17     | 1  | US-08-141-892A-8  |
| 26         | 17    | 89.5        | 17     | 1  | US-08-141-892A-14 |
| 27         | 17    | 89.5        | 17     | 1  | US-08-141-892A-41 |

|    |    |      |    |   |                   |                   |
|----|----|------|----|---|-------------------|-------------------|
| 28 | 17 | 89.5 | 17 | 2 | US-08-583-447A-8  | Sequence 8, Appli |
| 29 | 17 | 89.5 | 17 | 2 | US-08-583-447A-14 | Sequence 14, Appl |
| 30 | 17 | 89.5 | 17 | 2 | US-08-583-447A-41 | Sequence 41, Appl |
| 31 | 17 | 89.5 | 17 | 2 | US-08-467-920-8   | Sequence 8, Appli |
| 32 | 17 | 89.5 | 17 | 2 | US-08-467-920-14  | Sequence 14, Appl |
| 33 | 17 | 89.5 | 17 | 2 | US-08-467-920-41  | Sequence 41, Appl |
| 34 | 17 | 89.5 | 17 | 3 | US-08-635-930-8   | Sequence 8, Appli |
| 35 | 17 | 89.5 | 17 | 3 | US-08-635-930-14  | Sequence 14, Appl |
| 36 | 17 | 89.5 | 17 | 3 | US-08-635-930-41  | Sequence 41, Appl |
| 37 | 17 | 89.5 | 17 | 3 | US-09-193-997-8   | Sequence 8, Appli |
| 38 | 17 | 89.5 | 17 | 3 | US-09-193-997-14  | Sequence 14, Appl |
| 39 | 17 | 89.5 | 17 | 3 | US-09-193-997-41  | Sequence 41, Appl |
| 40 | 17 | 89.5 | 17 | 3 | US-09-138-237A-8  | Sequence 8, Appli |
| 41 | 17 | 89.5 | 17 | 3 | US-09-138-237A-14 | Sequence 14, Appl |
| 42 | 17 | 89.5 | 17 | 3 | US-09-138-237A-41 | Sequence 41, Appl |
| 43 | 16 | 84.2 | 16 | 1 | US-08-141-892A-9  | Sequence 9, Appli |
| 44 | 16 | 84.2 | 16 | 1 | US-08-141-892A-15 | Sequence 15, Appl |
| 45 | 16 | 84.2 | 16 | 2 | US-08-583-447A-9  | Sequence 9, Appli |

## ALIGNMENTS

RESULT 1  
US-08-141-892A-2  
; Sequence 2, Application US/08141892A  
; Patent No. 5518888  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: ST Receptor Binding Compounds and Methods  
; TITLE OF INVENTION: of Using the Same  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 55188888ris  
; STREET: One liberty Place - 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: U.S.A.  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch disk, 720 Kb  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/141,892A  
; FILING DATE: 26-OCT-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: TJU-0903  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-141-892A-2

Query Match 100.0%; Score 19; DB 1; Length 19;  
Best Local Similarity 100.0%; Pred. No. 8,9e-14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNTFYCCELCCNPACAGCY 19  
Db 1 NNTFYCCELCCNPACAGCY 19

us-10-775-481a-2.olg.ra

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,920-0
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,892
FILING DATE: 26-OCT-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-15899
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-467-920-2

Query Match 100.0%; Score
Best Local Similarity 100.0%; Pred:
Matches 19; Conservative 0; Mismatches 0

QY 1 NNTFYCCELCNCPACGY 19
Db 1 NNTFYCCELCNCPACGY 19

RESULT 4
US-08-635-930-2
Sequence 2, Application US/08635930
Patent No. 6,060,037
GENERAL INFORMATION:
APPLICANT: Waldman, Scott A.
TITLE OF INVENTION: Compositions Th
TITLE OF INVENTION: Colorectal Canc
TITLE OF INVENTION: The Same
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz
STREET: One Liberty Place, 46th Fl
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: WINDOWS 3.1
SOFTWARE: WordPerfect 6.0/6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/635,930-0
FILING DATE: 26-APR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,892
FILING DATE: 26-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/305,056
FILING DATE: 13-SEP-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark

```

```

RESULT 2
US-08-583-447A-2
; Sequence 2, Application US/08583447A
; Patent No. 5879656
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; TITLE OF INVENTION: ST Receptor Binding Compounds and
; TITLE OF INVENTION: Methods of Using the Same
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5879656ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: Wordperfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,447A
; FILING DATE: 05-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/141,892
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1702
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-583-447A-2

Query Match 100.0%; Score 19; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 8.9e-14;
Matches 19; Conservative 0; Mismatches 0; Indels 0;

QY 1 NNTFYCCCLCCNPACAGCY 19
Db 1 NNTFYCCCLCCNPACAGCY 19

RESULT 3
US-08-467-920-2
; Sequence 2, Application US/08467920
; Patent No. 5962220
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; TITLE OF INVENTION: Compositions That Specifically
; TITLE OF INVENTION: Bind to Colorectal Cancer Cells
; TITLE OF INVENTION: And Methods of Using the Same
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103

```

REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-1360  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-635-930-2

Query Match 100.0%; Score 19; DB 3; Length 19;  
Best Local Similarity 100.0%; Pred. No. 8.9e-14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNTFYCCELCNCPACGY 19  
Db 1 NNTFYCCELCNCPACGY 19

## RESULT 5

US-09-193-997-2  
Sequence 2, Application US/09193997  
Patent No. 6087109

GENERAL INFORMATION:  
APPLICANT: Waldman, Scott A.  
TITLE OF INVENTION: Compositions That Specifically  
Bind To Colorectal Cancer Cells  
TITLE OF INVENTION: And Methods Of Using The Same  
NUMBER OF SEQUENCES: 54  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
ADDRESSEE: No. 6087109ris  
STREET: One Liberty Place, 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/193,997

FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/467,920  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:

NAME: DeLuca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-1589  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-193-997-2

Query Match 100.0%; Score 19; DB 3; Length 19;  
Best Local Similarity 100.0%; Pred. No. 8.9e-14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNTFYCCELCNCPACGY 19

Db 1 NNTFYCCELCNCPACGY 19

## RESULT 6

US-09-138-237A-2  
Sequence 2, Application US/09138237A  
Patent No. 6268159  
GENERAL INFORMATION:  
APPLICANT: Waldman, Scott A.  
TITLE OF INVENTION: ST Receptor Binding Compounds and Methods  
TITLE OF INVENTION: Of Using the Same  
NUMBER OF SEQUENCES: 54  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 6268159ris  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19103

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch disk, 720 Kb  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/138,237A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/141,892  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:

NAME: DeLuca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-0903  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-138-237A-2

Query Match 100.0%; Score 19; DB 3; Length 19;  
Best Local Similarity 100.0%; Pred. No. 8.9e-14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNTFYCCELCNCPACGY 19  
Db 1 NNTFYCCELCNCPACGY 19

## RESULT 7

US-08-141-892A-7  
Sequence 7, Application US/08141892A  
Patent No. 5518888  
GENERAL INFORMATION:

APPLICANT: Waldman, Scott A.  
TITLE OF INVENTION: ST Receptor Binding Compounds and Methods  
TITLE OF INVENTION: Of Using the Same  
NUMBER OF SEQUENCES: 54  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 5518888ris  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19103  
COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch disk, 720 Kb  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/141,892A  
; FILING DATE: 26-OCT-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: TJU-0903  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-141-892A-7

Query Match 94.7%; Score 18; DB 1; Length 18;  
Best Local Similarity 100.0%; Pred. No. 9.6e-13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTFYCCCLCCNPACAGC 18  
|||||  
Db 1 NTFYCCCLCCNPACAGC 18

RESULT 8  
US-08-141-892A-13  
; Sequence 13, Application US/08141892A  
; Patent No. 5518888  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: ST Receptor Binding Compounds and Methods  
; TITLE OF INVENTION: of Using the Same  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 5518888ris  
; STREET: One Liberty Place - 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: U.S.A.  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch disk, 720 Kb  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/141,892A  
; FILING DATE: 26-OCT-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: TJU-0903  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids

; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-141-892A-13

Query Match 94.7%; Score 18; DB 1; Length 18;  
Best Local Similarity 100.0%; Pred. No. 9.6e-13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NTFYCCCLCCNPACAGCY 19  
|||||  
Db 1 NTFYCCCLCCNPACAGCY 18

RESULT 9  
US-08-141-892A-38  
; Sequence 38, Application US/08141892A  
; Patent No. 5518888  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: ST Receptor Binding Compounds and Methods  
; TITLE OF INVENTION: of Using the Same  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 5518888ris  
; STREET: One Liberty Place - 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: U.S.A.  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch disk, 720 Kb  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/141,892A  
; FILING DATE: 26-OCT-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: TJU-0903  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 38:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-141-892A-38

Query Match 94.7%; Score 18; DB 1; Length 18;  
Best Local Similarity 100.0%; Pred. No. 9.6e-13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NTFYCCCLCCNPACAGCY 19  
|||||  
Db 1 NTFYCCCLCCNPACAGCY 18

RESULT 10  
US-08-583-447A-7  
; Sequence 7, Application US/08583447A  
; Patent No. 5879656  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: ST Receptor Binding Compounds and



```
; TITLE OF INVENTION: Methods of Using the Same
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5879656ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,447A
; FILING DATE: 05-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/141,892
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1702
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-583-447A-7

Query Match          94.7%; Score 18; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 9.6e-13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NNTFYCCELCCNPACAGC 18
Db 1 NNTFYCCELCCNPACAGC 18

RESULT 11
US-08-583-447A-13
; Sequence 13, Application US/08583447A
; Patent No. 5879656
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; TITLE OF INVENTION: ST Receptor Binding Compounds and
; METHODS OF USING THE SAME
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5879656ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,447A
; FILING DATE: 05-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/141,892
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1702
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-583-447A-7

Query Match          94.7%; Score 18; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 9.6e-13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NNTFYCCELCCNPACAGC 18
Db 1 NNTFYCCELCCNPACAGC 18

RESULT 12
US-08-583-447A-13
; Sequence 13, Application US/08583447A
; Patent No. 5879656
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; TITLE OF INVENTION: ST Receptor Binding Compounds and
; METHODS OF USING THE SAME
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5879656ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,447A
; FILING DATE: 05-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/141,892
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1702
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-583-447A-7

Query Match          94.7%; Score 18; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 9.6e-13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NNTFYCCELCCNPACAGC 18
Db 1 NNTFYCCELCCNPACAGC 18
```

```
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1702
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-583-447A-13

Query Match          94.7%; Score 18; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 9.6e-13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 NTFYCCELCCNPACAGCY 19
Db 1 NTFYCCELCCNPACAGCY 18

RESULT 12
US-08-583-447A-38
; Sequence 38, Application US/08583447A
; Patent No. 5879656
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; TITLE OF INVENTION: ST Receptor Binding Compounds and
; METHODS OF USING THE SAME
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5879656ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,447A
; FILING DATE: 05-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/141,892
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1702
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-583-447A-38

Query Match          94.7%; Score 18; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 9.6e-13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 2 NTFYCCCLCCNPACAGCY 19  
|||||||  
Db 1 NTFYCCCLCCNPACAGCY 18

RESULT 13  
US-08-467-920-7  
; Sequence 7, Application US/08467920  
; Patent No. 5962220  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: Compositions That Specifically  
; TITLE OF INVENTION: Bind To Colorectal Cancer Cells  
; TITLE OF INVENTION: And Methods Of Using The Same  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; STREET: One Liberty Place, 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/467,920  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/141,892  
; FILING DATE: 26-OCT-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: TJU-1589  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-467-920-7  
Query Match 94.7%; Score 18; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 9.6e-13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
RESULT 14  
US-08-467-920-13  
; Sequence 13, Application US/08467920  
; Patent No. 5962220  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: Compositions That Specifically  
; TITLE OF INVENTION: Bind To Colorectal Cancer Cells  
; TITLE OF INVENTION: And Methods Of Using The Same  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &

; ADDRESSEE: No. 5962220ris  
; STREET: One Liberty Place, 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/467,920  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/141,892  
; FILING DATE: 26-OCT-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: TJU-1589  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-467-920-13  
Query Match 94.7%; Score 18; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 9.6e-13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
RESULT 15  
US-08-467-920-38  
; Sequence 38, Application US/08467920  
; Patent No. 5962220  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: Compositions That Specifically  
; TITLE OF INVENTION: Bind To Colorectal Cancer Cells  
; TITLE OF INVENTION: And Methods Of Using The Same  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; STREET: One Liberty Place, 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/467,920  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/141,892  
; FILING DATE: 26-OCT-1993

```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca; Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1589
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-467-920-38

```

```

Query Match      94.7%; Score 18; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 9.6e-13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 NTFYCCCLCCNPACAGCY 19
      |||||
Db      1 NTFYCCCLCCNPACAGCY 18

```

Search completed: March 26, 2005, 17:27:26  
 Job time : 25.7857 secs

*This Page Blank (uspio)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2005, 17:25:02 ; Search time 61.4107 Seconds  
(without alignments)  
102.440 Million cell updates/sec

Title: US-10-775-481A-2

Perfect score: 19

Sequence: 1 NNTFYCELCCNACAGCY 19

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1407402 seqs, 331100923 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1407402

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*

2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*

3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*

4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*

5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*

6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*

7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*

8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*

9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*

10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*

11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*

12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*

13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*

14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*

15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*

16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*

17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*

18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*

19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*

20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description      |
|------------|-------|-------------|--------|----|------------------|
| 1          | 19    | 100.0       | 19     | 15 | US-10-621-684-2  |
| 2          | 19    | 100.0       | 72     | 17 | US-10-796-719-20 |
| 3          | 18    | 94.7        | 18     | 15 | US-10-621-684-7  |
| 4          | 18    | 94.7        | 18     | 15 | US-10-621-684-13 |
| 5          | 18    | 94.7        | 18     | 15 | US-10-621-684-38 |
| 6          | 18    | 94.7        | 18     | 17 | US-10-796-719-2  |
| 7          | 18    | 94.7        | 18     | 17 | US-10-796-719-4  |
| 8          | 17    | 89.5        | 17     | 15 | US-10-621-684-8  |
| 9          | 17    | 89.5        | 17     | 15 | US-10-621-684-14 |
| 10         | 17    | 89.5        | 17     | 15 | US-10-621-684-41 |
| 11         | 16    | 84.2        | 16     | 15 | US-10-621-684-9  |
| 12         | 16    | 84.2        | 16     | 15 | US-10-621-684-15 |
| 13         | 15    | 78.9        | 15     | 15 | US-10-621-684-10 |
| 14         | 15    | 78.9        | 15     | 15 | US-10-621-684-16 |
| 15         | 15    | 78.9        | 15     | 15 | US-10-796-719-11 |
| 16         | 15    | 78.9        | 15     | 15 | US-10-621-684-42 |
| 17         | 14    | 73.7        | 14     | 15 | US-10-621-684-11 |
| 18         | 14    | 73.7        | 14     | 15 | US-10-621-684-17 |
| 19         | 14    | 73.7        | 14     | 15 | US-10-621-684-43 |
| 20         | 14    | 73.7        | 14     | 15 | US-10-621-684-44 |
| 21         | 13    | 68.4        | 13     | 15 | US-10-621-684-12 |
| 22         | 13    | 68.4        | 13     | 15 | US-10-621-684-45 |
| 23         | 11    | 57.9        | 11     | 15 | US-10-621-684-31 |
| 24         | 11    | 57.9        | 11     | 15 | US-10-621-684-52 |
| 25         | 11    | 57.9        | 11     | 15 | US-10-621-684-30 |
| 26         | 11    | 57.9        | 11     | 15 | US-10-621-684-36 |
| 27         | 11    | 57.9        | 11     | 15 | US-10-621-684-29 |
| 28         | 11    | 57.9        | 11     | 15 | US-10-621-684-35 |
| 29         | 11    | 57.9        | 11     | 15 | US-10-621-684-28 |
| 30         | 11    | 57.9        | 11     | 15 | US-10-621-684-34 |
| 31         | 11    | 57.9        | 11     | 15 | US-10-621-684-27 |
| 32         | 11    | 57.9        | 11     | 15 | US-10-621-684-33 |
| 33         | 11    | 57.9        | 11     | 15 | US-10-621-684-40 |
| 34         | 11    | 57.9        | 11     | 14 | US-10-107-814-23 |
| 35         | 11    | 57.9        | 11     | 15 | US-10-371-966-1  |
| 36         | 11    | 57.9        | 11     | 15 | US-10-371-966-2  |
| 37         | 11    | 57.9        | 11     | 15 | US-10-621-684-5  |
| 38         | 11    | 57.9        | 11     | 17 | US-10-796-719-1  |
| 39         | 11    | 57.9        | 11     | 17 | US-10-796-719-26 |
| 40         | 11    | 57.9        | 11     | 17 | US-10-796-719-27 |
| 41         | 11    | 57.9        | 11     | 17 | US-10-479-606-7  |
| 42         | 11    | 57.9        | 11     | 17 | US-10-796-719-39 |
| 43         | 11    | 57.9        | 11     | 17 | US-10-796-719-40 |
| 44         | 11    | 57.9        | 11     | 17 | US-10-796-719-21 |
| 45         | 10    | 52.6        | 13     | 15 | US-10-621-684-32 |

ALIGNMENTS

RESULT 1

US-10-621-684-2

Sequence 2, Application US/10621684

Publication No. US20040029182A1

GENERAL INFORMATION:

APPLICANT: Waldman, Scott A.

TITLE OF INVENTION: ST Receptor Binding Compounds and Methods of Using the Same

NUMBER OF SEQUENCES: 56

CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. US20040029182A1ris

STREET: One Liberty Place, 46th Floor

CITY: Philadelphia

STATE: Pennsylvania

COUNTRY: USA

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: Windows

SOFTWARE: Wordperfect 6.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/621.684

FILING DATE: 17-Jul-2003

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/583.447A

FILING DATE: 05-JAN-1996

APPLICATION NUMBER: US 08/141.892

FILING DATE: 26-OCT-1993

ATTORNEY/AGENT INFORMATION:

NAME: DeLuca, Mark

REGISTRATION NUMBER: 33,229

REFERENCE/DOCKET NUMBER: TJJ-1702

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-568-3100

```
;
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-621-684-2

Query Match 100.0%; Score 19; DB 15; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.1e-14;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNTFYCCCLCCNPACAGY 19
Db 1 NNTFYCCCLCCNPACAGY 19

RESULT 2
US-10-796-719-20
; Sequence 20, Application US/10796719
; Publication No. US20050020811A1
; GENERAL INFORMATION:
; APPLICANT: Currie, Mark G.
; APPLICANT: Mahajan-Miklos, Shalina
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TITLE OF INVENTION: TREATMENT OF GASTROINTESTINAL DISORDERS
; FILE REFERENCE: 14184-043001
; CURRENT APPLICATION NUMBER: US/10/796,719
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US 10/766,735
; PRIOR FILING DATE: 2004-01-28
; PRIOR APPLICATION NUMBER: US 60/443,098
; PRIOR FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 60/471,288
; PRIOR FILING DATE: 2003-05-15
; PRIOR APPLICATION NUMBER: US 60/519,460
; PRIOR FILING DATE: 2003-11-12
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; TYPE: PRT
; LENGTH: 72
; ORGANISM: Escherichia coli
US-10-796-719-20

Query Match 100.0%; Score 19; DB 17; Length 72;
Best Local Similarity 100.0%; Pred. No. 1.6e-13;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNTFYCCCLCCNPACAGY 19
Db 54 NNTFYCCCLCCNPACAGY 72

RESULT 3
US-10-621-684-7
; Sequence 7, Application US/10621684
; Publication No. US20040029182A1
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; TITLE OF INVENTION: ST Receptor Binding Compounds and
; TITLE OF INVENTION: Methods of Using the Same
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. US20040029182A1ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/621,684
; FILING DATE: 17-Jul-2003
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,447A
; FILING DATE: 05-JAN-1996
; APPLICATION NUMBER: US 08/141,892
; FILING DATE: 26-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1702
; TELECOMMUNICATION INFORMATION:
```

```
;
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/621,684
; FILING DATE: 17-Jul-2003
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,447A
; FILING DATE: 05-JAN-1996
; APPLICATION NUMBER: US 08/141,892
; FILING DATE: 26-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1702
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-621-684-7

Query Match 94.7%; Score 18; DB 15; Length 18;
Best Local Similarity 100.0%; Pred. No. 7.2e-13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNTFYCCCLCCNPACAGC 18
Db 1 NNTFYCCCLCCNPACAGC 18

RESULT 4
US-10-621-684-13
; Sequence 13, Application US/10621684
; Publication No. US20040029182A1
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; TITLE OF INVENTION: ST Receptor Binding Compounds and
; TITLE OF INVENTION: Methods of Using the Same
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. US20040029182A1ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/621,684
; FILING DATE: 17-Jul-2003
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,447A
; FILING DATE: 05-JAN-1996
; APPLICATION NUMBER: US 08/141,892
; FILING DATE: 26-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1702
; TELECOMMUNICATION INFORMATION:
```

; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:  
US-10-621-684-13

Query Match 94.7%; Score 18; DB 15; Length 18;  
Best Local Similarity 100.0%; Pred. No. 7.2e-13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NTFYCCCLCCNPACAGCY 19  
Db 1 NTFYCCCLCCNPACAGCY 18

## RESULT 5

US-10-621-684-38  
; Sequence 38, Application US/10621684  
; Publication No. US20040029182A1  
; GENERAL INFORMATION:

APPLICANT: Waldman, Scott A.  
TITLE OF INVENTION: ST Receptor Binding Compounds and  
Methods of Using the Same  
NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. US20040029182A1ris  
STREET: One Liberty Place, 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: Wordperfect 6.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/621.684  
FILING DATE: 17-Jul-2003  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/583.447A  
FILING DATE: 05-JAN-1996  
APPLICATION NUMBER: US 08/141.892  
FILING DATE: 26-OCT-1993

ATTORNEY/AGENT INFORMATION:

NAME: DeLuca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: T00-1702

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 38:

SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid

MOLECULE TYPE: linear

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 38:

US-10-621-684-38

Query Match 94.7%; Score 18; DB 15; Length 18;  
Best Local Similarity 100.0%; Pred. No. 7.2e-13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NTFYCCCLCCNPACAGCY 19  
Db 1 NTFYCCCLCCNPACAGCY 18

## RESULT 6

US-10-796-719-2  
; Sequence 2, Application US/10796719  
; Publication No. US20050020811A1  
; GENERAL INFORMATION:

APPLICANT: Currie, Mark G.  
APPLICANT: Mahajan-Miklos, Shalina

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE  
TREATMENT OF GASTROINTESTINAL DISORDERS  
FILE REFERENCE: 14184-043001

CURRENT APPLICATION NUMBER: US/10/796,719  
CURRENT FILING DATE: 2004-03-09

PRIOR APPLICATION NUMBER: US 10/766,735

PRIOR FILING DATE: 2004-01-28

PRIOR APPLICATION NUMBER: US 60/443,098

PRIOR FILING DATE: 2003-01-28

PRIOR APPLICATION NUMBER: US 60/471,288

PRIOR FILING DATE: 2003-05-15

PRIOR APPLICATION NUMBER: US 60/519,460

PRIOR FILING DATE: 2003-11-12

NUMBER OF SEQ ID NOS: 149

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 2

LENGTH: 18

TYPE: PRT

ORGANISM: Escherichia coli

US-10-796-719-2

Query Match 94.7%; Score 18; DB 17; Length 18;  
Best Local Similarity 100.0%; Pred. No. 7.2e-13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NTFYCCCLCCNPACAGCY 19  
Db 1 NTFYCCCLCCNPACAGCY 18

## RESULT 7

US-10-796-719-4  
; Sequence 4, Application US/10796719  
; Publication No. US20050020811A1  
; GENERAL INFORMATION:

APPLICANT: Currie, Mark G.  
APPLICANT: Mahajan-Miklos, Shalina

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE  
TREATMENT OF GASTROINTESTINAL DISORDERS  
FILE REFERENCE: 14184-043001

CURRENT APPLICATION NUMBER: US/10/796,719  
CURRENT FILING DATE: 2004-03-09

PRIOR APPLICATION NUMBER: US 10/766,735

PRIOR FILING DATE: 2004-01-28

PRIOR APPLICATION NUMBER: US 60/443,098

PRIOR FILING DATE: 2003-01-28

PRIOR APPLICATION NUMBER: US 60/471,288

PRIOR FILING DATE: 2003-05-15

PRIOR APPLICATION NUMBER: US 60/519,460

PRIOR FILING DATE: 2003-11-12

NUMBER OF SEQ ID NOS: 149

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 4

LENGTH: 18

TYPE: PRT

ORGANISM: Citrobacter freundii

US-10-796-719-4

Query Match 94.7%; Score 18; DB 17; Length 18;  
Best Local Similarity 100.0%; Pred. No. 7.2e-13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NTFYCCCLCCNPACAGCY 19  
Db 1 NTFYCCCLCCNPACAGCY 18

Db 1 NTFYCCCLCCNPACAGCY 18

## RESULT 8

US-10-621-684-8

; Sequence 8, Application US/10621684

; Publication No. US20040029182A1

; GENERAL INFORMATION:

; APPLICANT: Waldman, Scott A.

; TITLE OF INVENTION: ST Receptor Binding Compounds and

; METHODS OF USING THE SAME

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. US20040029182A1ris

; STREET: One Liberty Place, 46th Floor

; CITY: Philadelphia

; STATE: Pennsylvania

; COUNTRY: USA

; ZIP: 19103

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: Windows

; SOFTWARE: WordPerfect 6.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/621,684

; FILING DATE: 17-Jul-2003

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/583,447A

; FILING DATE: 05-JAN-1996

; APPLICATION NUMBER: US 08/141,892

; FILING DATE: 26-OCT-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Deluca, Mark

; REGISTRATION NUMBER: 33,229

; REFERENCE/DOCKET NUMBER: TJU-1702

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 215-568-3100

; TELEFAX: 215-568-3439

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 17 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; SEQUENCE DESCRIPTION: SEQ ID NO: 8:

US-10-621-684-8

## Query Match

Best Local Similarity 89.5%; Score 17; DB 15; Length 17;

Mismatches 0; Conservative 0; Indels 0; Gaps 0;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NTFYCCCLCCNPACAGC 18

Db 1 NTFYCCCLCCNPACAGC 17

## RESULT 9

US-10-621-684-14

; Sequence 14, Application US/10621684

; Publication No. US20040029182A1

; GENERAL INFORMATION:

; APPLICANT: Waldman, Scott A.

; TITLE OF INVENTION: ST Receptor Binding Compounds and

; METHODS OF USING THE SAME

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. US20040029182A1ris

; STREET: One Liberty Place, 46th Floor

; CITY: Philadelphia

; STATE: Pennsylvania

; COUNTRY: USA

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: Windows

SOFTWARE: WordPerfect 6.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/621,684

FILING DATE: 17-Jul-2003

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/583,447A

FILING DATE: 05-JAN-1996

APPLICATION NUMBER: US 08/141,892

FILING DATE: 26-OCT-1993

ATTORNEY/AGENT INFORMATION:

NAME: Deluca, Mark

REGISTRATION NUMBER: 33,229

REFERENCE/DOCKET NUMBER: TJU-1702

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-568-3100

TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 17 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 14:

US-10-621-684-14

## Query Match

Best Local Similarity 89.5%; Score 17; DB 15; Length 17;

Mismatches 0; Conservative 0; Indels 0; Gaps 0;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TFYCCCLCCNPACAGCY 19

Db 1 TFYCCCLCCNPACAGCY 17

## RESULT 10

US-10-621-684-41

; Sequence 41, Application US/10621684

; Publication No. US20040029182A1

; GENERAL INFORMATION:

; APPLICANT: Waldman, Scott A.

; TITLE OF INVENTION: ST Receptor Binding Compounds and

; METHODS OF USING THE SAME

; NUMBER OF SEQUENCES: 56

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. US20040029182A1ris

; STREET: One Liberty Place, 46th Floor

; CITY: Philadelphia

; STATE: Pennsylvania

; COUNTRY: USA

; ZIP: 19103

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: Windows

; SOFTWARE: WordPerfect 6.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/621,684

FILING DATE: 17-Jul-2003

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/583,447A

FILING DATE: 05-JAN-1996

APPLICATION NUMBER: US 08/141,892

FILING DATE: 26-OCT-1993

ATTORNEY/AGENT INFORMATION:

NAME: Deluca, Mark

REGISTRATION NUMBER: 33,229



REFERENCE/DOCKET NUMBER: TJU-1702  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 41:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 41:  
US-10-621-684-41

Query Match 89.5%; Score 17; DB 15; Length 17;  
Best Local Similarity 100.0%; Pred. No. 8.4e-12;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NTFYCCCLCCNPACAGC 18

Db 1 NTFYCCCLCCNPACAGC 17

## RESULT 11

US-10-621-684-9  
Sequence 9, Application US/10621684  
Publication No. US20040029182A1  
GENERAL INFORMATION:  
APPLICANT: Waldman, Scott A.  
TITLE OF INVENTION: ST Receptor Binding Compounds and  
Methods of Using the Same  
NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. US20040029182A1ris  
STREET: One Liberty Place, 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: WordPerfect 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/621.684  
FILING DATE: 17-Jul-2003  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/583.447A  
FILING DATE: 05-JAN-1996  
REFERENCE/DOCKET NUMBER: TJU-1702  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 9:  
US-10-621-684-9

Query Match 84.2%; Score 16; DB 15; Length 16;  
Best Local Similarity 100.0%; Pred. No. 9.9e-11;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TFYCCCLCCNPACAGC 18

Db 1 TFYCCCLCCNPACAGC 16

## RESULT 12

US-10-621-684-15  
Sequence 15, Application US/10621684  
Publication No. US20040029182A1  
GENERAL INFORMATION:  
APPLICANT: Waldman, Scott A.  
TITLE OF INVENTION: ST Receptor Binding Compounds and  
Methods of Using the Same  
NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. US20040029182A1ris  
STREET: One Liberty Place, 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: WordPerfect 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/621.684  
FILING DATE: 17-Jul-2003  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/583.447A  
FILING DATE: 05-JAN-1996  
REFERENCE/DOCKET NUMBER: TJU-1702  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 15:  
US-10-621-684-15

Query Match 84.2%; Score 16; DB 15; Length 16;  
Best Local Similarity 100.0%; Pred. No. 9.9e-11;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 FYCCCLCCNPACAGCY 19

Db 1 FYCCCLCCNPACAGCY 16

## RESULT 13

US-10-621-684-10  
Sequence 10, Application US/10621684  
Publication No. US20040029182A1  
GENERAL INFORMATION:  
APPLICANT: Waldman, Scott A.  
TITLE OF INVENTION: ST Receptor Binding Compounds and  
Methods of Using the Same  
NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. US20040029182A1ris  
STREET: One Liberty Place, 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: WordPerfect 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/621.684  
FILING DATE: 17-Jul-2003  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/583.447A  
FILING DATE: 05-JAN-1996  
REFERENCE/DOCKET NUMBER: TJU-1702  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-10-621-684-10

Query Match 84.2%; Score 16; DB 15; Length 16;  
Best Local Similarity 100.0%; Pred. No. 9.9e-11;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TFYCCCLCCNPACAGC 18

```
/
/ COUNTRY: USA
/ ZIP: 19103
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: Windows
/ SOFTWARE: WordPerfect 6.1
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/621,684
/ FILING DATE: 17-Jul-2003
/ CLASSIFICATION: 435
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/583,447A
/ FILING DATE: 05-JAN-1996
/ APPLICATION NUMBER: US 08/141,892
/ FILING DATE: 26-OCT-1993
/
/ ATTORNEY/AGENT INFORMATION:
/ NAME: DeLuca, Mark
/ REFERENCE/DOCKET NUMBER: TJU-1702
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 215-568-3100
/ TELEFAX: 215-568-3439
/
/ INFORMATION FOR SEQ ID NO: 10:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 15 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-621-684-10

Query Match 78.9%; Score 15; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 FYCCELCNCPACAGC 18
Db 1 FYCCELCNCPACAGC 15

RESULT 14
US-10-621-684-16
; Sequence 16, Application US/10621684
; Publication No. US20040029182A1
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; TITLE OF INVENTION: ST Receptor Binding Compounds and
; Methods of Using the Same
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. US20040029182A1ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: WordPerfect 6.1
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/621,684
; FILING DATE: 17-Jul-2003
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,447A
; FILING DATE: 05-JAN-1996
; APPLICATION NUMBER: US 08/141,892
; FILING DATE: 26-OCT-1993
;
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca, Mark
; REFERENCE/DOCKET NUMBER: TJU-1702
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
;
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 42:
US-10-621-684-42

Query Match 78.9%; Score 15; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 YCCELCNCPACAGY 19
Db 1 YCCELCNCPACAGY 15

RESULT 15
US-10-621-684-42
; Sequence 42, Application US/10621684
; Publication No. US20040029182A1
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; TITLE OF INVENTION: ST Receptor Binding Compounds and
; Methods of Using the Same
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. US20040029182A1ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: WordPerfect 6.1
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/621,684
; FILING DATE: 17-Jul-2003
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,447A
; FILING DATE: 05-JAN-1996
; APPLICATION NUMBER: US 08/141,892
; FILING DATE: 26-OCT-1993
;
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca, Mark
; REFERENCE/DOCKET NUMBER: TJU-1702
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
;
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 42:
US-10-621-684-42

Query Match 78.9%; Score 15; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 5 YCCELCNPNACAGCY 19  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1 YCCELCNPNACAGCY 15

Search completed: March 26, 2005, 17:44:30  
Job time : 61.4107 secs

*This Page Blank (uspio)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2005, 17:10:31 ; Search time 19.6786 Seconds  
(without alignments)  
92.899 Million cell updates/sec

Title: US-10-775-481A-2

Perfect score: 19

Sequence: 1 NNTFYCCELCNCPACGY 19

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR\_79:.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 19    | 100.0       | 72     | 1 QHEC1  | heat-stable entero |
| 2          | 18    | 94.7        | 18     | 2 A50103 | heat-stable entero |
| 3          | 11    | 57.9        | 72     | 1 QHEC4  | heat-stable entero |
| 4          | 11    | 57.9        | 72     | 1 QHECIB | heat-stable entero |
| 5          | 10    | 52.6        | 18     | 1 QHEC2  | heat-stable entero |
| 6          | 10    | 52.6        | 53     | 2 S68705 | heat-stable entero |
| 7          | 9     | 47.4        | 66     | 2 S31652 | enterotoxin - Vers |
| 8          | 9     | 47.4        | 71     | 2 S25659 | heat-stable entero |
| 9          | 6     | 31.6        | 13     | 2 A28953 | alpha-conotoxin SI |
| 10         | 6     | 31.6        | 15     | 1 NTKNAG | alpha-conotoxin GI |
| 11         | 6     | 31.6        | 17     | 2 A45334 | heat-stable entero |
| 12         | 6     | 31.6        | 19     | 2 A44379 | alpha-conotoxin SI |
| 13         | 6     | 31.6        | 65     | 2 S34671 | heat-stable entero |
| 14         | 6     | 31.6        | 78     | 1 QHVC1  | heat-stable entero |
| 15         | 6     | 31.6        | 90     | 2 D85845 | probable lysis pro |
| 16         | 6     | 31.6        | 620    | 2 H82761 | sulfur deprivation |
| 17         | 6     | 31.6        | 924    | 2 E71476 | alanine-tRNA ligas |
| 18         | 5     | 26.3        | 55     | 2 S14212 | NADH2 dehydrogenas |
| 19         | 5     | 26.3        | 75     | 2 H84631 | hypothetical prote |
| 20         | 5     | 26.3        | 108    | 2 A71043 | hypothetical prote |
| 21         | 5     | 26.3        | 110    | 2 I74319 | gene EHS-2 protein |
| 22         | 5     | 26.3        | 111    | 2 B72476 | hypothetical prote |
| 23         | 5     | 26.3        | 116    | 2 S75378 | hypothetical prote |
| 24         | 5     | 26.3        | 124    | 2 C96762 | hypothetical prote |
| 25         | 5     | 26.3        | 129    | 2 A24255 | chorion class A pr |
| 26         | 5     | 26.3        | 132    | 2 A10543 | probable secreted  |
| 27         | 5     | 26.3        | 133    | 2 T17300 | hypothetical prote |
| 28         | 5     | 26.3        | 135    | 2 S34815 | nifU protein - Rh  |
| 29         | 5     | 26.3        | 139    | 1 Q28P87 | gene 2.8 protein - |

RESULT 1

QHEC1

heat-stable enterotoxin ST-I precursor - Escherichia coli

N/Alternate names: heat-stable enterotoxin estal

C/Species: Escherichia coli

C/Date: 31-Aug-1980 #sequence revision 31-Aug-1980 #text\_change 09-Jul-2004

C/Accession: A01822; A30985; A36732; J0374; I51932

R/So, M.; McCarthy, B.J.

Proc. Natl. Acad. Sci. U.S.A. 77, 4011-4015, 1980

A/Title: Nucleotide sequence of the bacterial transposon Tn1681 encoding a heat-stable (S

A/Reference number: A01822; MUID:81054703; PMID:6254008

A/Accession: A01822

A/Molecule type: DNA

A/Residues: 1-72 <LAZ>

A/Cross-references: UNIPROT:P01559; GB:V00612; GB:J01831; NID:G43704; PIDN:CAA23883.1; P

R/Lazure, C.; Seidah, N.G.; Chretien, M.; Lallier, R.; St-Pierre, S.

Can. J. Biochem. Cell Biol. 61, 287-292, 1983

A/Title: Primary structure determination of Escherichia coli heat-stable enterotoxin of

A/Reference number: A30985; MUID:83284515; PMID:6349752

A/Accession: A30985

A/Molecule type: protein

A/Residues: 55-72 <LAZ2>

A/Experimental source: strain F11

R/Dallas, W.S.

J. Bacteriol. 172, 5490-5493, 1990

A/Title: The heat-stable toxin I gene from Escherichia coli 18D.

A/Reference number: A36732; MUID:90368614; PMID:2203756

A/Accession: A36732

A/Molecule type: DNA

A/Residues: 1-72 <DAL>

A/Cross-references: GB:M58746; NID:G145860; PIDN:AAA62776.1; PID:G145861

A/Experimental source: strain 18D

R/Stieglitz, H.; Cervantes, L.; Robledo, R.; Fonseca, R.; Covarrubias, L.; Bolivar, F.;

Plasmid 20, 42-53, 1988

A/Title: Cloning, sequencing, and expression in ficoll-generated minicells of an Escheri

A/Reference number: J0373; MUID:89202548; PMID:3071819

A/Accession: J0374

A/Molecule type: DNA

A/Residues: 1-72 <STI>

R/Sekizaki, T.; Akashi, H.; Terakado, N.

Am. J. Vet. Res. 46, 909-912, 1985

A/Title: Nucleotide sequences of the genes for Escherichia coli heat-stable enterotoxin

A/Reference number: I51932; MUID:85249571; PMID:2990268

A/Accession: I51932

A/Status: translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-69, 'p', 71-72 <RES>

A/Cross-references: GB:M25607; NID:G147877; PIDN:AAA24653.1; PID:G147878

C/Comment: Both heat-stable and heat-labile enterotoxins are produced by pathogenic stra

ular sizes.

C/Superfamily: heat-stable enterotoxin ST

C/Keywords: enterotoxin; heat-stable protein

heterodisulfide re  
hypothetical prote  
hypothetical prote  
probable rRNA meth  
BioY family protei  
conserved hypotet  
probable pentose-5  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
aquaporin homolog  
protein ZC123.3 [I  
cob(I)alamin adeno

F;1-19/Domain: signal sequence #status predicted <SIG>  
 F;20-54/Domain: propeptide #status predicted <PRO>  
 F;55-72/Product: heat-stable enterotoxin ST-I #status experimental <MAT>  
 F;59-64,60-68,63-71/Disulfide bonds: #status predicted

Query Match 100.0%; Score 19; DB 1; Length 72;  
 Best Local Similarity 100.0%; Pred. No. 2e-15;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNTFYCCELCCNPACAGCY 19  
 |||||  
 Db 54 NNTFYCCELCCNPACAGCY 72

## RESULT 2

A60103 heat-stable enterotoxin ST-Ia - Citrobacter freundii

C;Species: Citrobacter freundii  
 C;Date: 10-Nov-1992 #sequence\_revision 10-Nov-1992 #text\_change 09-Jul-2004  
 C;Accession: A60103  
 R;Guarino, A.; Giannella, R.; Thompson, M.R.

Infect. Immun. 57, 649-652, 1989  
 A;Title: Citrobacter freundii produces an 18-amino-acid heat-stable enterotoxin identical to that of Escherichia coli  
 A;Reference number: A60103; MUID:89108617; PMID:2912902

A;Accession: A60103  
 A;Molecule type: protein  
 A;Residues: 1-18 <GUA>  
 A;Cross-references: UNIPROT:Q7M0U3  
 C;Superfamily: heat-stable enterotoxin ST

Query Match 94.7%; Score 18; DB 2; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-14;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NTFYCCELCCNPACAGCY 19  
 |||||  
 Db 1 NTFYCCELCCNPACAGCY 18

## RESULT 3

QHEC4 heat-stable enterotoxin STA4 precursor - Escherichia coli

C;Species: Escherichia coli  
 C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 09-Jul-2004  
 C;Accession: J0373; A35978  
 R;Stieglicz, H.; Cervantes, L.; Robledo, R.; Fonseca, R.; Covarrubias, L.; Bolivar, F.;

Plasmid 20, 42-53, 1988  
 A;Title: Cloning, sequencing, and expression in ficoll-generated minicells of an Escherichia coli heat-stable enterotoxin ST-I  
 A;Reference number: J0373; MUID:89202548; PMID:3071819

A;Accession: J0373  
 A;Molecule type: DNA  
 A;Residues: 1-72 <STI>

C;Cross-references: UNIPROT:P07965; GB:J03311; NID:gl47875; PIDN:AAA24652.1; PID:gl47876  
 R;Zhou, X.; Shen, L.P.; Chi, C.W.  
 Toxin 28, 453-456, 1990

A;Title: Isolation and nucleotide sequence determination of a gene encoding a heat-stable enterotoxin ST-I from Escherichia coli  
 A;Reference number: A35978; MUID:90273381; PMID:2190361

A;Accession: A35978  
 A;Molecule type: DNA  
 A;Residues: 1-72 <ZHO>

C;Genetics:

A;Gene: estA4  
 C;Superfamily: heat-stable enterotoxin ST  
 C;Keywords: enterotoxin; heat-stable protein  
 F;1-19/Domain: signal sequence #status predicted <SIG>  
 F;20-53/Domain: propeptide #status predicted <PRO>  
 F;54-72/Product: heat-stable enterotoxin #status predicted <MAT>

F;59-64,60-68,63-71/Disulfide bonds: #status predicted  
 Query Match 57.9%; Score 11; DB 1; Length 72;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-06;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 YCCELCCNPAC 15  
 |||||  
 Db 58 YCCELCCNPAC 68

## RESULT 4

QHEC1B

heat-stable enterotoxin ST-Ib precursor - Escherichia coli  
 N;Alternate names: heat-stable enterotoxin ST-A2  
 C;Species: Escherichia coli

C;Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 09-Jul-2004  
 C;Accession: JS0292; A33068; A33067; A30567  
 R;Moseley, S.L.; Hardy, J.W.; Hug, M.I.; Echeverria, P.; Falkow, S.  
 Infect. Immun. 39, 1167-1174, 1983

A;Title: Isolation and nucleotide determination of a gene encoding a heat-stable enterotoxin ST-Ib from Escherichia coli  
 A;Reference number: JS0292; MUID:83184648; PMID:6341230

A;Accession: JS0292

A;Molecule type: DNA

A;Residues: 1-72 <MOS>

A;Cross-references: UNIPROT:Q47185; UNIPROT:P07965; GB:M34916; NID:gl46407; PIDN:AAA2399  
 R;Dwarakanath, P.; Viewswarlah, S.S.; Subrahmanyam, Y.V.B.K.; Shanthi, G.; Jagannatha, I.  
 Gene 81, 219-226, 1989

A;Title: Cloning and hyperexpression of a gene encoding the heat-stable toxin of Escherichia coli  
 A;Reference number: A33068; MUID:90034194; PMID:2680769

A;Accession: A33068

A;Molecule type: DNA

A;Residues: 1-18, 'A', 20-72 <DWA>

A;Cross-references: GB:M23255; NID:gl48029; PIDN:AAA24686.1; PID:gl48030

A;Note: The authors translated the codon AAG for residue 2 as Val and CTA for residue 34 as Leu. The authors also noted that the codon AAG for residue 2 as Val and CTA for residue 34 as Leu. The authors also noted that the codon AAG for residue 2 as Val and CTA for residue 34 as Leu.  
 R;Aimoto, S.; Takao, T.; Shimonishi, Y.; Hara, S.; Takeda, T.; Takeda, Y.; Miwatani, T.  
 Eur. J. Biochem. 129, 257-263, 1982

A;Title: Amino acid sequence of heat-stable enterotoxin produced by human enterotoxigenic Escherichia coli  
 A;Reference number: A33067; MUID:83105138; PMID:6759126

A;Accession: A33067

A;Molecule type: protein

A;Residues: 54-72 <ALM>

R;Guzman-Verduzco, L.M.; Kupersztosch, Y.M.

Infect. Immun. 57, 645-648, 1989

A;Title: Rectification of two Escherichia coli heat-stable enterotoxin allele sequences  
 A;Reference number: A30567; MUID:89108616; PMID:2643580

A;Accession: A30567

A;Molecule type: DNA

A;Residues: 1-18, 'A', 20-24, 'AG', 27-41, 'V', 43-44, 'N', 46, 'E', 48, 'S', 50-72 <GUZ>

A;Cross-references: GB:M18345; NID:gl45862; PIDN:AAA23729.1; PID:gl45863

C;Comment: This is one of the type I heat-stable enterotoxins that are methanol-soluble.

C;Genetics:

A;Gene: st

C;Superfamily: heat-stable enterotoxin ST

C;Keywords: enterotoxin; heat-stable protein

F;1-53/Domain: signal sequence and propeptide #status predicted <SIG>

F;54-72/Product: heat-stable enterotoxin ST-Ib #status experimental <MAT>

F;59-64,60-68,63-71/Disulfide bonds: #status experimental

Query Match 57.9%; Score 11; DB 1; Length 72;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-06;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 YCCELCCNPAC 15  
 |||||  
 Db 58 YCCELCCNPAC 68

## RESULT 5

QHEC2

heat-stable enterotoxin ST-2 - Escherichia coli

C;Species: Escherichia coli

C;Date: 06-Jul-1982 #sequence\_revision 06-Jul-1982 #text\_change 09-Jul-2004

C;Accession: A01823

R;Chan, S.K.; Giannella, R.A.

J. Biol. Chem. 256, 7744-7746, 1981

A;Title: Amino acid sequence of heat-stable enterotoxin produced by Escherichia coli

A;Reference number: A01823; MUID:81264141; PMID:7021541

A;Accession: A01823

A:Molecule type: protein  
A:Residues: 1-18 <CHA>  
A:Cross-references: UNIPROT:P01560  
A:Experimental source: strain 18D, serotype 0.42:k86:H37  
C:Comment: This enterotoxin is one of several, of differing molecular sizes, produced by isolates of the heat-stable enterotoxin ST-1.  
C:Superfamily: heat-stable enterotoxin ST  
C:Keywords: enterotoxin; heat-stable protein  
F:1-18/Product: heat-stable enterotoxin ST-2 #status experimental <MAT>  
F:5-10,6-14,9-17/Disulfide bonds: #status predicted

Query Match 52.6%; Score 10; DB 1; Length 18;  
Best Local Similarity 100.0%; Pred. No. 2.5e-05;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NTFYCCELC 11  
| | | | | | | |  
Db 1 NTFYCCELC 10

RESULT 6  
S68705  
heat-stable enterotoxin Y-STc - Yersinia enterocolitica  
C:Species: Yersinia enterocolitica  
C:Date: 25-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 13-Mar-1998  
C:Accession: S68705  
R:Yoshino, K.; Takao, T.; Huang, X.; Murata, H.; Nakao, H.; Takeda, T.; Shimonishi, Y.  
FEBS Lett. 362, 319-322, 1995  
A:Title: Characterization of a highly toxic, large molecular size heat-stable enterotoxin  
A:Reference number: S68705; MUID:95246844; PMID:7729521  
A:Accession: S68705  
A:Molecule type: protein  
A:Residues: 1-53 <YOS>  
A:Experimental source: strain 86-11  
C:Superfamily: heat-stable enterotoxin ST  
C:Keywords: enterotoxin; heat-stable protein  
F:41-46,42-50,45-53/Disulfide bonds: #status predicted

Query Match 52.6%; Score 10; DB 2; Length 53;  
Best Local Similarity 100.0%; Pred. No. 5.3e-05;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 CCELCNCPAC 15  
| | | | | | | |  
Db 41 CCELCNCPAC 50

RESULT 7  
S31652  
enterotoxin - Yersinia kristensenii  
C:Species: Yersinia kristensenii  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004  
C:Accession: S31652  
R:Ibrahim, A.; Liesack, W.; Stackebrandt, E.  
submitted to the EMBL Data Library, November 1992  
A:Reference number: S31652  
A:Accession: S31652  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-66 <IBR>  
A:Cross-references: UNIPROT:P31518; EMBL:X69218; NID:G48617; PIDN:CAA49152.1; PID:G48618  
C:Superfamily: heat-stable enterotoxin ST

Query Match 47.4%; Score 9; DB 2; Length 66;  
Best Local Similarity 100.0%; Pred. No. 0.0009;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 CCNPACAGC 18  
| | | | | | | |  
Db 58 CCNPACAGC 66

RESULT 8

heat-stable enterotoxin yst precursor - Yersinia enterocolitica  
C:Species: Yersinia enterocolitica  
C:Date: 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 09-Jul-2004  
C:Accession: S25659; A41474; A23114; S65849  
R:Ibrahim, A.; Liesack, W.; Pike, S.; Stackebrandt, E.  
FEMS Microbiol. Lett. 97, 63-66, 1992  
A:Title: The polymerase chain reaction: an epidemiological tool to differentiate between F:1-18/Product: heat-stable enterotoxin ST-2 #status experimental <MAT>  
A:Reference number: S25659  
A:Accession: S25659  
A:Status: preliminary; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-71 <IBR>  
A:Cross-references: UNIPROT:P07593; EMBL:X65999; NID:G48611; PIDN:CAA46801.1; PID:G48612  
R:Delor, I.; Kaeckenbeeck, A.; Wauters, G.; Cornelis, G.R.  
Infect. Immun. 58, 2983-2988, 1990  
A:Title: Nucleotide sequence of yst, the Yersinia enterocolitica gene encoding the heat-stable enterotoxin  
A:Reference number: A41474; MUID:90354067; PMID:2201642  
A:Accession: A41474  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-47, 'S', 49-71 <DEL>  
A:Cross-references: GB:U09235; NID:G487394; PIDN:AAA18472.1; PID:G487395  
R:Takao, T.; Tomimaga, N.; Yoshimura, S.; Shimonishi, Y.; Hara, S.; Inoue, T.; Miyama, A.  
Eur. J. Biochem. 152, 199-206, 1985  
A:Title: Isolation, primary structure and synthesis of heat-stable enterotoxin produced by Yersinia enterocolitica  
A:Reference number: A23114; MUID:86004705; PMID:4043080  
A:Accession: A23114  
A:Molecule type: protein  
A:Residues: 54-71 <TAK>  
R:Milulek, A.V.; Delor, I.; Ha Thi, V.; Cornelis, G.R.  
Mol. Microbiol. 14, 905-915, 1994  
A:Title: Regulation of the Yersinia enterocolitica enterotoxin yst gene. Influence of growth conditions on yst expression  
A:Reference number: S65849; MUID:95231297; PMID:7715452  
A:Accession: S65849  
A:Status: preliminary; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-43 <MIK>  
A:Cross-references: EMBL:U09235  
C:Genetics:  
A:Gene: yst  
C:Superfamily: heat-stable enterotoxin ST  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-41/Domain: propeptide #status predicted <PRO>  
F:42-71/Product: heat-stable enterotoxin yst #status predicted <MAT>

Query Match 47.4%; Score 9; DB 2; Length 71;  
Best Local Similarity 100.0%; Pred. No. 0.00095;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 CCNPACAGC 18  
| | | | | | | |  
Db 63 CCNPACAGC 71

RESULT 9  
A28953  
alpha-conotoxin SI - cone shell (Conus striatus)  
C:Species: Conus striatus (striated cone)  
C:Date: 30-Jun-1989 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
C:Accession: A28953  
R:Zafaralla, G.C.; Ramilo, C.; Gray, W.R.; Karlstrom, R.; Olivera, B.M.; Cruz, L.J.  
Biochemistry 27, 7102-7105, 1988  
A:Title: Phylogenetic specificity of cholinergic ligands: alpha-conotoxin SI.  
A:Reference number: A28953; MUID:89062448; PMID:3196703  
A:Accession: A28953  
A:Molecule type: protein  
A:Residues: 1-13 <ZAP>  
A:Cross-references: UNIPROT:P15471  
A:Note: this sequence was confirmed by chemical synthesis  
C:Comment: This paralytic toxin from a fish-hunting cone snail inhibits the acetylcholine receptor  
C:Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; postsynaptic neurotoxin

F;2-7,3-13/Disulfide bonds: #status experimental  
F;13/Modified site: amidated carboxyl end (Cys) #status experimental

Query Match 31.6%; Score 6; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.96;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 CCNPAC 15  
Db |||||

Db 2 CCNPAC 7

## RESULT 10

NTKNAG

alpha-conotoxin GIA [validated] - cone shell (Conus geographus)

N;Alternate names: alpha-Crx-GIA

N;Contains: alpha-conotoxin GI

C;Species: Conus geographus (geography cone)

C;Date: 24-Sep-1981 #sequence\_revision 24-Sep-1981 #text\_change 09-Jul-2004

C;Accession: A01782

R;Gray, W.R.; Luque, A.; Olivera, B.M.; Barrett, J.; Cruz, L.J.

J. Biol. Chem. 256, 4734-4740, 1981

A;Title: Peptide toxins from Conus geographus venom.

A;Reference number: A92320; MUID:81191854; PMID:7014556

A;Accession: A01782

A;Molecule type: protein

A;Residues: 1-15 <GRA>

A;Cross-references: UNIPROT:P01519

R;Gray, W.R.; Rivier, J.E.; Galyean, R.; Cruz, L.J.; Olivera, B.M.

J. Biol. Chem. 258, 12247-12251, 1983

A;Title: Conotoxin M1. Disulfide bonding and conformational states.

A;Reference number: A92396; MUID:84032400; PMID:6630187

A;Contents: annotation; disulfide bonds

R;Guddat, L.W.; Shan, L.; Martin, J.L.; Edmundson, A.B.; Gray, W.R.

submitted to the Brookhaven Protein Data Bank, May 1996

A;Reference number: A66253; PDB:1NOT

A;Contents: annotation; X-ray crystallography, 1.2 angstroms, residues 1-13

R;Guddat, L.W.; Martin, J.A.; Shan, L.; Edmundson, A.B.; Gray, W.R.

Biochemistry 35, 11329-11335, 1996

A;Title: Three-dimensional structure of the alpha-conotoxin GI at 1.2 angstroms resolution

A;Reference number: A58592; MUID:96378624; PMID:8784187

A;Contents: annotation; X-ray crystallography, 1.2 angstroms

R;Fardi, A.; Galdes, A.; Florance, J.; Manicotte, D.

Biochemistry 28, 5494-5501, 1989

A;Title: Solution structures of alpha-conotoxin GI determined by two-dimensional NMR spectroscopy

A;Reference number: A30629; MUID:89375269; PMID:2775719

A;Contents: annotation; conformation by (1)H-NMR

C;Comment: This alpha-conotoxin, as an acetylcholine receptor inhibitor, is a postsynaptic

C;Superfamily: alpha-conotoxin

C;Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; blocked carboxyl end

F;1-13/Product: conotoxin GI #status experimental <GIA>

F;2-7,3-13/Disulfide bonds: #link GIA #status predicted

F;2-7,3-13/Disulfide bonds: #link GIA #status experimental

F;13/Modified site: amidated carboxyl end (Cys) (amide in mature form from following gly

F;15/Modified site: blocked carboxyl end (Lys) (probably amidated) #status experimental

Query Match 31.6%; Score 6; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 CCNPAC 15

Db |||||

Db 2 CCNPAC 7

## RESULT 11

A54534

heat-stable enterotoxin - Vibrio mimicus (fragment)

C;Species: Vibrio mimicus

C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 03-May-1996

C;Accession: A54534

R;Arita, M.; Honda, T.; Miwatani, T.; Takeda, T.; Takao, T.; Shimonishi, Y.

FEBS Microbiol. Lett. 79, 105-110, 1991

A;Title: Purification and characterization of a heat-stable enterotoxin of Vibrio mimicus  
A;Reference number: A54534

A;Accession: A54534

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-17 <ARI>

C;Superfamily: heat-stable enterotoxin ST

Query Match 31.6%; Score 6; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 CCNPAC 15

Db |||||

Db 7 CCNPAC 12

## RESULT 12

A44379

alpha-conotoxin SII - cone shell (Conus striatus)

C;Species: Conus striatus (striated cone)

C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 16-Jul-1999

C;Accession: A44379

R;Ramilo, C.A.; Zafaralla, G.C.; Nadasdi, L.; Hammerland, L.G.; Yoshikami, D.; Gray, W.R.

Biochemistry 31, 9919-9926, 1992

A;Title: Novel alpha- and omega-conotoxins from Conus striatus venom.

A;Reference number: A44379; MUID:93003172; PMID:1390774

A;Accession: A44379

A;Molecule type: protein

A;Residues: 1-19 <RAM>

A;Cross-references: CAS:143294-31-9; PIDN:AB23762.1; PID:9257934

A;Experimental source: venom

A;Note: sequence extracted from NCBI backbone (NCBIP:116000); structure confirmed by chen

C;Comment: This peptide is an acetylcholine receptor blocker.

C;Superfamily: alpha-conotoxin

C;Keywords: acetylcholine receptor inhibitor; postsynaptic neurotoxin; venom

F;2-18,3-8,4-14/Disulfide bonds: #status predicted

F;19/Modified site: amidated carboxyl end (Ser) #status absent

Query Match 31.6%; Score 6; DB 2; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 CCNPAC 15

Db |||||

Db 3 CCNPAC 8

## RESULT 13

S34671

heat-stable enterotoxin - Vibrio cholerae (fragment)

C;Species: Vibrio cholerae

C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004

C;Accession: S34671

R;Rosolini, G.M.; Lombardi, G.; Guglielmetti, P.

submitted to the EMBL Data Library, July 1993

A;Description: Detection of the Vibrio cholerae heat-stable enterotoxin gene by polymerase

A;Reference number: S34671

A;Accession: S34671

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-65 <ROS>

A;Cross-references: UNIPROT:Q56643; EMBL:X74108

C;Superfamily: heat-stable enterotoxin ST

Query Match 31.6%; Score 6; DB 2; Length 65;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 CCNPAC 15

Db |||||

Db 59 CCNPAC 64



## RESULT 14

QHVCI  
heat-stable enterotoxin ST precursor - Vibrio cholerae  
C:Species: Vibrio cholerae  
C:Date: 17-Mar-1987 #sequence revision 30-Jun-1992 #text change 09-Jul-2004  
C:Accession: A1469; A01824; S34464; S34466; S34465; S34463  
R:Ogawa, A.; Kato, J.I.; Watanabe, H.; Nair, B.G.; Takeda, T.  
Infect. Immun. 58, 3325-3329, 1990  
A:Title: Cloning and nucleotide sequence of a heat-stable enterotoxin gene from Vibrio cholerae  
A:Reference number: A1469; MUID:90382953; PMID:2205577  
A:Accession: A1469  
A:Molecule type: DNA  
A:Residues: 1-78 <OGA>  
A:Cross-references: UNIPROT:P04429; GB:M85198; NID:g155237; PIDN:AAA64889.1;  
R:Takao, T.; Shimonishi, Y.; Kobayashi, M.; Nishimura, O.; Arita, M.; Takeda, T.; Honda,  
PEBS Lett. 193, 250-254, 1985  
A:Title: Amino acid sequence of heat-stable enterotoxin produced by Vibrio cholerae non-  
O1  
A:Reference number: A01824; MUID:86056320; PMID:4065341  
A:Accession: A01824  
A:Molecule type: protein  
A:Residues: 62-78 <TAK>  
A:Experimental source: non-O1 aerovar  
R:Yoshino, K.; Miyachi, M.; Takao, T.; Bag, P.K.; Xiaozhe, H.; Nair, G.B.; Takeda, T.; S  
PEBS Lett. 326, 83-86, 1993  
A:Title: Purification and sequence determination of heat-stable enterotoxin elaborated b  
y Vibrio cholerae non-O1  
A:Reference number: S34463; MUID:93314823; PMID:8325391  
A:Accession: S34464  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 61-78 <Y03>  
A:Accession: S34466  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 51-78 <Y05>  
A:Accession: S34465  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 60-78 <Y02>  
A:Accession: S34463  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 62-78 <Y04>  
C:Superfamily: enterotoxin; heat-stable enterotoxin ST  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-61/Domain: propeptide #status predicted <PRO>  
F:62-78/Product: heat-stable enterotoxin ST #status experimental <MAT>  
F:64-69,65-73,68-76/Disulfide bonds: #status predicted

Query Match 31.6%; Score 6; DB 1; Length 78;  
Best Local Similarity 100.0%; Pred.No.3.3;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 CCNPAC 15  
Db 68 CCNPAC 73

## RESULT 15

D85845  
probable lysis protein S of prophage CP-933V Z3340 [imported] - Escherichia coli (strain  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence revision 16-Feb-2001 #text change 09-Jul-2004  
C:Accession: D85845  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouisis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: D85845  
A>Status: preliminary

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2005, 16:51:20 ; Search time 74.9821 Seconds  
(without alignments)  
129.758 Million cell updates/sec

Title: US-10-775-481A-2

Perfect score: 19

Sequence: 1 NNTFYCCELCNPAAGCY 19

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID        | Description        |
|------------|-------|-------------|--------|--------------|--------------------|
| 1          | 19    | 100.0       | 72     | 1 HST1_ECOLI | P01559 escherichia |
| 2          | 18    | 94.7        | 72     | 2 Q7M0U3     | Q7m0u3 citrobacter |
| 3          | 16    | 84.2        | 61     | 2 Q5VEG9     | Q5veg9 escherichia |
| 4          | 11    | 57.9        | 61     | 2 Q5VEG7     | Q5veg7 escherichia |
| 5          | 11    | 57.9        | 61     | 2 Q5VEG8     | Q5veg8 escherichia |
| 6          | 11    | 57.9        | 72     | 1 HST2_ECOLI | Q47185 escherichia |
| 7          | 11    | 57.9        | 72     | 1 HST3_ECOLI | P07965 escherichia |
| 8          | 10    | 52.6        | 72     | 1 HSTB_ECOLI | P01560 escherichia |
| 9          | 10    | 52.6        | 72     | 1 HSTC_YEREN | O50319 yersinia en |
| 10         | 9     | 47.4        | 66     | 1 HST_YERKR  | P1518 yersinia kr  |
| 11         | 9     | 47.4        | 71     | 1 HSTB_YEREN | P07593 yersinia en |
| 12         | 9     | 47.4        | 71     | 1 HSTB_YEREN | P74977 yersinia en |
| 13         | 6     | 31.6        | 65     | 1 CXAA_CONGE | P01519 conus geogr |
| 14         | 6     | 31.6        | 17     | 2 Q5R581     | Q5r581 vibrio chol |
| 15         | 6     | 31.6        | 18     | 2 Q5R580     | Q5r580 vibrio chol |
| 16         | 6     | 31.6        | 19     | 2 Q5R579     | Q5r579 vibrio chol |
| 17         | 6     | 31.6        | 28     | 2 Q5R578     | Q5r578 vibrio chol |
| 18         | 6     | 31.6        | 64     | 1 CXAI_CONST | P15471 conus stria |
| 19         | 6     | 31.6        | 66     | 2 Q56643     | Q56643 vibrio chol |
| 20         | 6     | 31.6        | 72     | 1 CXA2_CONST | P28879 conus stria |
| 21         | 6     | 31.6        | 78     | 1 HSTN_VIBCH | P04429 vibrio chol |
| 22         | 6     | 31.6        | 78     | 1 HSTO_VIBCH | Q07425 vibrio chol |
| 23         | 6     | 31.6        | 70     | 2 Q93G01     | Q93g01 vibrio mimi |
| 24         | 6     | 31.6        | 90     | 2 Q8X4M8     | Q8x4m8 escherichia |
| 25         | 6     | 31.6        | 96     | 2 Q7Y222     | Q7y222 stx1 conver |
| 26         | 6     | 31.6        | 96     | 2 Q9T1L3     | Q9t1l3 bacterioph  |
| 27         | 6     | 31.6        | 120    | 2 Q858Z4     | Q858z4 bacterioph  |
| 28         | 6     | 31.6        | 158    | 2 Q8SC52     | Q8sc52 stx2 conver |
| 29         | 6     | 31.6        | 276    | 2 Q96H89     | Q96h89 homo sapien |
| 30         | 6     | 31.6        | 415    | 2 Q8CZV0     | Q8czv0 yersinia pe |
| 31         | 6     | 31.6        | 556    | 2 Q8VUI3     | Q8vui3 desulfovibr |

32 6 31.6 619 2 Q9GNI9 Q9gni9-caenorhabdi  
33 6 31.6 620 2 Q87AF0 Q87af0 xyella fas  
34 6 31.6 620 2 Q9PF93 Q9pf93 xyella fas  
35 6 31.6 645 2 Q8A2H6 Q8a2h6 bacteroides  
36 6 31.6 3401 2 Q8IBQ3 Q8ibq3 plasmodium  
37 5 26.3 15 1 GUAN\_DIDMA P55936 didelphis m  
38 5 26.3 40 1 CXAC\_CONGE Q8erb2 conus geogr  
39 5 26.3 45 2 Q83AK0 Q83ak0 coxiella bu  
40 5 26.3 50 2 Q83BF3 Q83bf3 coxiella bu  
41 5 26.3 57 2 Q83A72 Q83a72 coxiella bu  
42 5 26.3 63 2 Q9STC3 Q9setc3 elaeis guin  
43 5 26.3 65 2 Q62MM8 Q62mm8 burkholderi  
44 5 26.3 66 2 Q6C751 Q6c751 yarrowia li  
45 5 26.3 66 2 Q688G0 Q688g0 oryza sativ

#### ALIGNMENTS

RESULT 1  
HST1\_ECOLI STANDARD; PRT; 72 AA.  
AC P01559; Q47653;  
DT 21-JUN-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Heat-stable enterotoxin ST-IA/ST-P precursor.  
GN Name=stal;  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TRANSPOSON=Tn1681;  
RX MEDLINE=81054703; PubMed=6254008;  
RA So M., McCarthy B.J.;  
RT "Nucleotide sequence of the bacterial transposon Tn1681 encoding a heat-stable (ST) toxin and its identification in enterotoxigenic Escherichia coli strains";  
RL Proc. Natl. Acad. Sci. U.S.A. 77:4011-4015(1980).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O42:K86:H37 / 18D / ETEC;  
RX MEDLINE=90368614; PubMed=2203756;  
RA Dallas W.S.;  
RT "The heat-stable toxin I gene from Escherichia coli 18D.";  
RL J. Bacteriol. 172:5490-5493(1990).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=85249571; PubMed=2990268;  
RA Sekizaki T., Akashi H., Terakado N.;  
RT "Nucleotide sequences of the genes for Escherichia coli heat-stable enterotoxin I of bovine, avian, and porcine origins.";  
RL Am. J. Vet. Res. 46:909-912(1985).  
RN [4]  
RP DISULFIDE BONDS.  
RX MEDLINE=87191003; PubMed=3552731; DOI=10.1016/0014-5793(87)80134-5;  
RA Shimonishi Y., Hidaka Y., Koizumi M., Hane M., Aimoto S., Takeda T.,  
RX Miwatani T., Takeda Y.;  
RT "Mode of disulfide bond formation of a heat-stable enterotoxin (STh) produced by a human strain of enterotoxigenic Escherichia coli.";  
RL FEBS Lett. 215:165-170(1987).  
RN [5]  
RP PROCESSING.  
RX MEDLINE=90368584; PubMed=2203746;  
RA Okamoto K., Takahara M.;  
RT "Synthesis of Escherichia coli heat-stable enterotoxin STh as a pre-pro form and role of the pro sequence in secretion.";  
RL J. Bacteriol. 172:5260-5265(1990).  
RN [6]  
RP X-RAY CRYSTALLOGRAPHY (0.89 ANGSTROMS) OF 59-71.  
RX MEDLINE=94312375; PubMed=8038153;

RA Sato T., Ozaki H., Kitagawa Y., Katsube Y., Shimonishi Y.;  
RT "Structural characteristics for biological activity of heat-stable  
RT enterotoxin produced by enterotoxigenic Escherichia coli: X-ray  
RT crystallography of weakly toxic and nontoxic analogs.";  
RL Biochemistry 33:18641-18650(1994).  
CC -1- FUNCTION: Toxin which activates the particulate form of guanylate  
CC cyclase and increases cyclic GMP levels within the host intestinal  
CC epithelial cells.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- SIMILARITY: Belongs to the heat-stable enterotoxin family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; V00612; CAA23883.1; -;  
DR EMBL; M58746; AAA62776.1; -;  
DR EMBL; M25607; AAA24653.1; -;  
DR PIR; A01822; QHEC1.  
DR PDB; 1ETL; X-ray; @=59-71.  
DR PDB; 1ETW; X-ray; @=-.  
DR PDB; 1ETN; X-ray; @=-.  
DR InterPro: IPR001489; Enterotoxin HS.  
DR Pfam: PF02048; Enterotoxin HS; 1.  
DR PROSITE; PS00273; ENTEROTOXIN\_H\_STABLE; 1.  
KW 3D-structure; Enterotoxin; Signal; Toxin; Transposable element.  
FT SIGNAL 1 19  
FT PROPEP 20 54  
FT PEPTIDE 55 72 Heat-stable enterotoxin ST-1A/ST-P.  
FT DISULFID 59 64  
FT DISULFID 60 68  
FT DISULFID 63 71  
FT CONFLICT 70 70 G -> P (in Ref. 3).  
FT TURN 61 62  
FT TURN 65 67  
FT TURN 69 70  
SQ SEQUENCE 72 AA; 8075 MW; 9288B766B3988264 CRC64;  
  
Query Match 100.0%; Score 19; DB 1; Length 72;  
Best Local Similarity 100.0%; Pred. No. 8.8e-15;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 NNTFYCCELCNCPACAGCY 19  
Db 54 NNTFYCCELCNCPACAGCY 72  
|||||  
  
RESULT 2  
ID Q7M0U3 PRELIMINARY; PRT; 18 AA.  
AC Q7M0U3;  
DT 01-MAR-2004 (T-EMBLrel. 26, Created)  
DT 01-MAR-2004 (T-EMBLrel. 26, Last sequence update)  
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)  
DE Heat-stable enterotoxin ST-1a.  
OS Citrobacter freundii.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Citrobacter.  
OX NCBI\_TaxID=546;  
RN [1]  
RP MEDLINE=89108617; PubMed=2912902;  
RX Guarino A., Giannella R., Thompson M.R.;  
RA "Citrobacter freundii produces an 18-amino-acid heat-stable  
RT enterotoxin identical to the 18-amino-acid Escherichia coli heat-  
RT stable enterotoxin (ST 1a)".  
RL Infect. Immun. 57:649-652(1989).  
DR PIR; A60103.  
DR HSSP; P01559; 1ETN.

DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0009405; P:pathogenesis; IEA.  
DR InterPro: IPR001489; Enterotoxin HS.  
DR Pfam; PF02048; Enterotoxin HS; 1.  
DR PROSITE; PS00273; ENTEROTOXIN\_H\_STABLE; 1.  
SQ SEQUENCE 18 AA; 1978 MW; D4D975F49D60064F CRC64;  
  
Query Match 94.7%; Score 18; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 4.7e-14;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 NTFYCCCELCNCPACAGCY 19  
Db 1 NTFYCCCELCNCPACAGCY 18  
|||||  
  
RESULT 3  
ID Q6VEG9 PRELIMINARY; PRT; 61 AA.  
AC Q6VEG9;  
DT 05-JUL-2004 (T-EMBLrel. 27, Created)  
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)  
DE Heat-stable enterotoxin ST-1a (Fragment).  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=F7682;  
RX PubMed=15364995;  
RA Reischl U., Youssef M.T., Wolf H., Hyttia-Trees E., Strockbine N.A.;  
RT "Real-time fluorescence PCR assays for detection and characterization  
RT of heat-labile I and heat-stable I enterotoxin genes from  
RT enterotoxigenic Escherichia coli.";  
RL J. Clin. Microbiol. 42:4092-4100(2004).  
DR EMBL; AY342057; AAQ92974.1; -;  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0009405; P:pathogenesis; IEA.  
DR InterPro: IPR001489; Enterotoxin HS.  
DR Pfam; PF02048; Enterotoxin HS; 1.  
FT NON\_TER 1 1  
SQ SEQUENCE 61 AA; 6927 MW; 646D4AE2F899D957 CRC64;  
  
Query Match 84.2%; Score 16; DB 2; Length 61;  
Best Local Similarity 100.0%; Pred. No. 2.6e-11;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 NNTFYCCELCNCPACA 16  
Db 43 NNTFYCCELCNCPACA 58  
|||||  
  
RESULT 4  
ID Q6VEG7 PRELIMINARY; PRT; 61 AA.  
AC Q6VEG7;  
DT 05-JUL-2004 (T-EMBLrel. 27, Created)  
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)  
DE Heat-stable enterotoxin ST 1b (Fragment).  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=B544;  
RX PubMed=15364995;  
RA Reischl U., Youssef M.T., Wolf H., Hyttia-Trees E., Strockbine N.A.;  
RT "Real-time fluorescence PCR assays for detection and characterization  
RT of heat-labile I and heat-stable I enterotoxin genes from

```

RT enterotoxigenic Escherichia coli.;
RL J. Clin. Microbiol. 42:4092-4100 (2004).
DR EMBL; AY342059; AAQ92976.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR001489; Enterotoxin_HS.
DR Pfam; PF02048; Enterotoxin_HS; 1.
DR PROSITE; PS00273; ENTEROTOXIN_H_STABLE; 1.
FT NON_TER 1
SQ SEQUENCE 61 AA; 6556 MW; 89788D3FAB3DCA0A CRC64;

Query Match 57.9%; Score 11; DB 2; Length 61;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 YCCELCCNPAC 15
Db 47 YCCELCCNPAC 57

RESULT 5
Q6VEG8
ID Q6VEG8 PRELIMINARY; PRT; 61 AA.
AC Q6VEG8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Heat-stable enterotoxin ST Ib (Fragment).
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C4046;
RA Reischl U., Yousef M.T., Wolf H., Hyttia-Trees E., Stroockbine N.A.;
RX "Real-time fluorescence PCR assays for detection and characterization
RT of heat-labile I and heat-stable I enterotoxin genes from
RT enterotoxigenic Escherichia coli.";
RL J. Clin. Microbiol. 42:4092-4100 (2004).
DR EMBL; AY342058; AAQ92975.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR001489; Enterotoxin_HS.
DR Pfam; PF02048; Enterotoxin_HS; 1.
DR PROSITE; PS00273; ENTEROTOXIN_H_STABLE; 1.
FT NON_TER 1
SQ SEQUENCE 61 AA; 6658 MW; 1D75955D7AF0DED2 CRC64;

Query Match 57.9%; Score 11; DB 2; Length 61;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 YCCELCCNPAC 15
Db 47 YCCELCCNPAC 57

RESULT 6
HST2_ECOLI
ID HST2_ECOLI STANDARD; PRT; 72 AA.
AC Q47185;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Heat-stable enterotoxin A2 precursor (STA2).
GN Name=sta2;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=89108616; PubMed=2643580;
RA Guxman-Verduzio L.M., Kupersztosch Y.M.;
RT "Rectification of two Escherichia coli heat-stable enterotoxin allele
RT sequences and lack of biological effect of changing the carboxy-
RT terminal tyrosine to histidine.";
RL Infect. Immun. 57:645-648 (1989).
CC -1- FUNCTION: Toxin which activates the particulate form of guanylate
CC cyclase and increases cyclic GMP levels within the host intestinal
CC epithelial cells.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the heat-stable enterotoxin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M18345; AAA23729.1; -.
DR PIR; JS0292; QHECIB.
DR HSSP; P01559; 1ETN.
DR InterPro; IPR001489; Enterotoxin_HS.
DR Pfam; PF02048; Enterotoxin_HS; 1.
DR PROSITE; PS00273; ENTEROTOXIN_H_STABLE; 1.
KW Enterotoxin; Signal; Toxin.
FT SIGNAL 1 19 Potential.
FT PROPEP 20 53 By similarity.
FT PEPTIDE 54 72 Heat-stable enterotoxin A2.
FT DISULFID 59 64 By similarity.
FT DISULFID 60 68 By similarity.
FT DISULFID 63 71 By similarity.
SQ SEQUENCE 72 AA; 7895 MW; D87850306E06B260 CRC64;

Query Match 57.9%; Score 11; DB 1; Length 72;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 YCCELCCNPAC 15
Db 58 YCCELCCNPAC 68

RESULT 7
HST3_ECOLI
ID HST3_ECOLI STANDARD; PRT; 72 AA.
AC P07965; P26588;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Heat-stable enterotoxin A3/A4 precursor (STA3/STA4) (ST-H) (ST-H).
GN Name=sta3; Synonyms=sta4;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89202548; PubMed=3071819;
RA Stieglitz H., Kupersztosch Y.M.;
RA Bolivar F., Kupersztosch Y.M.;
RT "Cloning, sequencing, and expression in Ficolli-generated minicells of
RT an Escherichia coli heat-stable enterotoxin gene.";
RL Plasmid 20:42-53 (1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=153837-2;
RX MEDLINE=83184648; PubMed=6341230;
RA Moseley S.L., Hardy J.W., Huq M.I., Echeverria P., Falkow S.;
RT "Isolation and nucleotide sequence determination of a gene encoding a
RT heat-stable enterotoxin of Escherichia coli.";

```



```
AC O50319;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Heat-stable enterotoxin C precursor (Y-57C).
GN Name=y57C;
OS Yersinia enterocolitica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=630;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=86-11;
RX MEDLINE=97202479; PubMed=9049998; DOI=10.1006/mpat.1996.0094;
RA Huang X., Yoshino K., Nakao H., Takeda T.;
RT "Nucleotide sequence of a gene encoding the novel Yersinia
RT enterocolitica heat-stable enterotoxin that includes a pro-region-like
RT sequence in its mature toxin molecule.";
RL Microb. Pathog. 22:89-97(1997).
RN [2]
RP SEQUENCE OF 20-72.
RC STRAIN=Serotype O:3;
RX MEDLINE=95246844; PubMed=7729521; DOI=10.1016/0014-5793(95)00267-D;
RA Yoshino K., Takeda T., Huang X., Murata H., Nakao H., Takeda T.,
RA Shimonishi Y.;
RT "Characterization of a highly toxic, large molecular size heat-stable
RT enterotoxin produced by a clinical isolate of Yersinia
RT enterocolitica.";
RL FEBS Lett. 362:319-322(1995).
CC -!- FUNCTION: Toxin which activates the particulate form of guanylate
CC cyclase and increases cyclic GMP levels within the host intestinal
CC epithelial cells. Highly toxic.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the heat-stable enterotoxin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D63578; BAA23656.1; -
DR HSSP; P01559; 1ETN.
DR InterPro; IPR001489; Enterotoxin HS.
DR Pfam; PF02048; Enterotoxin HS_1
DR PROSITE; PS00273; ENTEROTOXIN_H_STABLE; 1.
KW Direct protein sequencing; Enterotoxin; Signal; Toxin.
FT SIGNAL 1 19
FT CHAIN 20 72 Heat-stable enterotoxin C.
FT DISULFID 60 65 By similarity.
FT DISULFID 61 69 By similarity.
FT DISULFID 64 72 By similarity.
SQ SEQUENCE 72 AA; 7639 MW; 7C0D83893C2F981D CRC64;

Query Match 52.6%; Score 10; DB 1; Length 72;
Best Local Similarity 100.0%; Pred. No. 0.00032;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 6 CCELCNCPAC 15
Db 60 CCELCNCPAC 69

RESULT 10
HST_YERKR
ID HST_YERKR STANDARD; PRT; 66 AA.
AC P31518;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Heat-stable enterotoxin precursor.
```

```
GN Name=y57;
OS Yersinia kristensenii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=28152;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IP490 / Serotype O:12,25;
RA Ibrahim A., Liesack W., Stackebrandt E.;
RL Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Toxin which activates the particulate form of guanylate
CC cyclase and increases cyclic GMP levels within the host intestinal
CC epithelial cells (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the heat-stable enterotoxin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X69218; CAA49152.1; -
DR PIR; S31652; S31652.
DR HSSP; P01559; 1ETN.
DR InterPro; IPR001489; Enterotoxin HS.
DR Pfam; PF02048; Enterotoxin HS_1
DR PROSITE; PS00273; ENTEROTOXIN_H_STABLE; 1.
KW Enterotoxin; Signal.
FT SIGNAL 1 19 Potential.
FT CHAIN 20 50 Heat-stable enterotoxin (By similarity).
FT DISULFID 51 56 By similarity.
FT DISULFID 54 59 By similarity.
FT DISULFID 55 63 By similarity.
FT DISULFID 58 66 By similarity.
SQ SEQUENCE 66 AA; 7068 MW; 27BE7006675CC075 CRC64;

Query Match 47.4%; Score 9; DB 1; Length 66;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 10 CCNPACAGC 18
Db 58 CCNPACAGC 66

RESULT 11
HSTA_YEREN
ID HSTA_YEREN STANDARD; PRT; 71 AA.
AC P07593;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Heat-stable enterotoxin A precursor (YST-A).
GN Name=y57A; Synonyms=y57;
OS Yersinia enterocolitica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=630;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Serotype O:8;
RX MEDLINE=93051243; PubMed=1427005; DOI=10.1016/0378-1097(92)90364-T;
RA Ibrahim A., Liesack W., Pike S., Stackebrandt E.;
RT "The polymerase chain reaction: an epidemiological tool to
RT differentiate between two clusters of pathogenic Yersinia
RT enterocolitica strains.";
RL FEMS Microbiol. Lett. 76:63-66(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=W1024 / Serotype O:9;
```

RX MEDLINE=90354067; PubMed=2201642;  
 RA Delor I., Kaeckenbeck A., Wauters G., Cornelis G.R.;  
 RT "Nucleotide sequence of yst, the yersinia enterocolitica gene encoding  
 the heat-stable enterotoxin, and prevalence of the gene among  
 pathogenic and nonpathogenic yersiniae.";   
 RL Infect. Immun. 58:2983-2988(1990).  
 RN [3]  
 RP SEQUENCE OF 54-71.  
 RX MEDLINE=86004705; PubMed=4043080;  
 RA Takao T., Tomimaga N., Yoshimura S., Shimonishi Y., Hara S., Inoue T.,  
 RA Miyama A.;  
 RT "Isolation, primary structure and synthesis of heat-stable enterotoxin  
 produced by Yersinia enterocolitica.";   
 RL Eur. J. Biochem. 152:199-206(1985)  
 CC -1- FUNCTION: Toxin which activates the particulate form of guanylate  
 cyclase and increases cyclic GMP levels within the host intestinal  
 epithelial cells.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- INDUCTION: In cultured cells, expressed only at temperatures <30  
 degrees Celsius. Under conditions of high osmolality and alkaline  
 pH (as it is the case in the host's intestine), it is expressed at  
 37 degrees Celsius.  
 CC -1- SIMILARITY: Belongs to the heat-stable enterotoxin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X65999; CAA46801.1; -;  
 DR EMBL; U09235; AAA18472.1; -;  
 DR PIR; S25659; S25659.  
 DR HSSP; P01559; IETN.  
 DR InterPro; IPR001489; Enterotoxin HS.  
 DR Pfam; PF02048; Enterotoxin HS; 1\_  
 DR PROSITE; PS00273; ENTEROTOXIN\_H\_STABLE; 1.  
 KW Direct protein sequencing; Enterotoxin; Signal; Toxin.  
 FT SIGNAL 1 19 Potential.  
 FT PROPEP 20 53 Heat-stable enterotoxin A.  
 FT CHAIN 54 71 By similarity.  
 FT DISULFID 59 64 By similarity.  
 FT DISULFID 60 68 By similarity.  
 FT DISULFID 63 71 By similarity.  
 FT VARIANT 48 48 L -> S.  
 SQ SEQUENCE 71 AA; 7494 MW; 022F99B3800C861B CRC64;  
 Query Match 47.4%; Score 9; DB 1; Length 71;  
 Best Local Similarity 100.0%; Pred. No. 0.0047;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 10 CCNPACAGC 18  
 Db 63 CCNPACAGC 71  
 |||||  
 RESULT 12  
 HSTB\_YEREN STANDARD; PRT; 71 AA.  
 AC P7497;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Heat-stable enterotoxin B precursor (Y-STB).  
 GN Names=YSTB;  
 OS Yersinia enterocolitica.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Yersinia.  
 OX NCBI\_TaxID=630;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC STRAIN=84-50 / Serotype O:5;  
 RX MEDLINE=98008784; PubMed=9344780; DOI=10.1006/mpat.1997.0146;  
 RA Ramamurthy T., Yoshino K., Huang X., Nair G.B., Carniel E.,  
 RA Maruyama T., Fukushima H., Takeda T.;  
 RT "The novel heat-stable enterotoxin subtype gene (ystB) of Yersinia  
 enterocolitica: nucleotide sequence and distribution of the yst  
 genes.";   
 RL Microb. Pathog. 23:189-200(1997).  
 CC -1- FUNCTION: Toxin which activates the particulate form of guanylate  
 cyclase and increases cyclic GMP levels within the host intestinal  
 epithelial cells. Could play an important role in pathogenesis.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Belongs to the heat-stable enterotoxin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D88145; BAA13544.1; -;  
 DR HSSP; P01559; IETN.  
 DR InterPro; IPR001489; Enterotoxin HS.  
 DR Pfam; PF02048; Enterotoxin HS; 1\_  
 DR PROSITE; PS00273; ENTEROTOXIN\_H\_STABLE; 1.  
 KW Direct protein sequencing; Enterotoxin; Signal; Toxin.  
 FT SIGNAL 1 19 Potential.  
 FT PROPEP 20 52 Heat-stable enterotoxin B.  
 FT CHAIN 53 71 By similarity.  
 FT DISULFID 59 64 By similarity.  
 FT DISULFID 60 68 By similarity.  
 FT DISULFID 63 71 By similarity.  
 SQ SEQUENCE 71 AA; 7670 MW; ED6E9F61ACDD4F50 CRC64;  
 Query Match 47.4%; Score 9; DB 1; Length 71;  
 Best Local Similarity 100.0%; Pred. No. 0.0047;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 10 CCNPACAGC 18  
 Db 63 CCNPACAGC 71  
 |||||  
 RESULT 13  
 CXAA\_CONGE STANDARD; PRT; 15 AA.  
 ID CXAA\_CONGE  
 AC P01519;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Alpha-conotoxin GIA [contains: Alpha-conotoxin GI (G1)].  
 OS Conus geographus (Geography cone).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;  
 OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;  
 OC Neogastropoda; Conoidea; Conidae; Conus.  
 OX NCBI\_TaxID=6491;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=81191854; PubMed=7014556;  
 RA Gray W.R., Luque F.A., Olivera B.M., Barrett J., Cruz L.J.;  
 RT "Peptide toxins from Conus geographus venom.";   
 RL J. Biol. Chem. 256:4734-4740(1981).  
 RN [2]  
 RP DISULFIDE BONDS OF GI, AND SYNTHESIS OF GI.  
 RX MEDLINE=83105694; PubMed=7152021; DOI=10.1016/0014-5793(82)80820-X;  
 RA Nishiuchi Y., Sakakibara S.;  
 RT "Primary and secondary structure of conotoxin GI, a neurotoxic  
 tridecapeptide from a marine snail.";   
 RL FEBS Lett. 148:260-262(1982).  
 RN [3]  
 RP DISULFIDE BONDS OF GI, AND SYNTHESIS OF GI.



RX MEDLINE=84280842; PubMed=6466616;  
RA Gray W.R., Luque F.A., Galyean R., Atherton E., Sheppard R.C.,  
RA Stone B.L., Reyes A., Alford J., McIntosh M., Olivera B.M., Cruz L.J.,  
RA Rivier J.;  
RT "Conotoxin GI: disulfide bridges, synthesis, and preparation of  
RT iodinated derivatives.";  
RL Biochemistry 23:2796-2802 (1984).  
RN [4]  
RP COMPARISON WITH ALPHA-CONOTOXIN SI AND ALPHA-CONOTOXIN MI.  
RX MEDLINE=95034849; PubMed=7947815;  
RA Hann R.M., Pagan O.R., Eterovic V.A.;  
RT "The alpha-conotoxins GI and MI distinguish between the nicotinic  
RT acetylcholine receptor agonist sites while SI does not.";  
RL Biochemistry 33:14058-14063 (1994).  
RN [5]  
RP PHARMACOLOGICAL CHARACTERIZATION ON MOUSE MUSCLE-DERIVED BC3H-1 CELLS  
RP AND TORPEDO ELECTRIC ORGAN  
RX MEDLINE=95349531; PubMed=7623764;  
RA Groebe D.R., Dumm J.M., Levitan E.S., Abramson S.N.;  
RT "alpha-Conotoxins selectively inhibit one of the two acetylcholine  
RT binding sites of nicotinic receptors.";  
RL Mol. Pharmacol. 48:105-111 (1995).  
RN [6]  
RP MUTAGENESIS OF ARG-9.  
RX MEDLINE=97317090; PubMed=9174364; DOI=10.1021/bi970195w;  
RA Groebe D.R., Gray W.R., Abramson S.N.;  
RT "Determinants involved in the affinity of alpha-conotoxins GI and SI  
RT for the muscle subtype of nicotinic acetylcholine receptors.";  
RL Biochemistry 36:8469-8474 (1997).  
RN [7]  
RP X-RAY CRYSTALLOGRAPHY (1.2 ANGSTROMS) OF GI.  
RX MEDLINE=96378624; PubMed=8784187; DOI=10.1021/bi960820h;  
RA Guddat L.W., Martin J.A., Shan L., Edmundson A.B., Gray W.R.;  
RT "Three-dimensional structure of the alpha-conotoxin GI at 1.2-A  
RT resolution.";  
RL Biochemistry 35:11329-11335 (1996).  
RN [8]  
RP STRUCTURE BY NMR OF GI.  
RX MEDLINE=89352562; PubMed=2765514;  
RA Kobayashi Y., Ohkubo T., Kyogoku Y., Nishiuchi Y., Sakakibara S.,  
RA Braun W., Go N.;  
RT "Solution conformation of conotoxin GI determined by 1H nuclear  
RT magnetic resonance spectroscopy and distance geometry calculations.";  
RL Biochemistry 28:4853-4860 (1989).  
RN [9]  
RP STRUCTURE BY NMR OF GI.  
RX MEDLINE=89375269; PubMed=2775719;  
RA Pardi A., Gaides A., Fiorance J., Manicote D.;  
RT "Solution structures of alpha-conotoxin GI determined by two-  
RT dimensional NMR spectroscopy.";  
RL Biochemistry 28:5494-5501 (1989).  
RN [10]  
RP STRUCTURE BY NMR OF GI.  
RX MEDLINE=98321613; PubMed=9660176;  
RA Maslennikov I.V., Sobol A.G., Gladky K.V., Lugovskoy A.A.,  
RA Ostrovsky A.G., Tselin V.I., Ivanov V.T., Arseniev A.S.;  
RT "Two distinct structures of alpha-conotoxin GI in aqueous solution.";  
RL Eur. J. Biochem. 254:238-247 (1998).  
RN [11]  
RP STRUCTURE BY NMR OF GI.  
RX MEDLINE=98239743; PubMed=9571060; DOI=10.1006/jmbi.1998.1701;  
RA Gehrmann J., Alewood P.F., Craik D.J.;  
RT "Structure determination of the three disulfide bond isomers of alpha-  
RT conotoxin GI: a model for the role of disulfide bonds in structural  
RT stability.";  
RL J. Mol. Biol. 278:401-415 (1998).  
RN [12]  
RP STRUCTURE BY NMR OF AN ANTITOXIC ANALOG OF GI.  
RX MEDLINE=99438341; PubMed=10508392; DOI=10.1021/bi990558n;  
RA Mok K.H., Han K.H.;  
RT "NMR solution conformation of an antitoxic analogue of alpha-conotoxin  
RT GI: identification of a common nicotinic acetylcholine receptor  
RT alpha(1)-subunit binding surface for small ligands and alpha-

RT conotoxins.";  
RL Biochemistry 38:11895-11904 (1999).  
CC -1- FUNCTION: Alpha-conotoxins act on postsynaptic membranes, they  
CC bind to the nicotinic acetylcholine receptors (nAChR) and thus  
CC inhibit them. The higher affinity site for alpha-conotoxin GI is  
CC the alpha/delta site on mouse muscle-derived BC3H-1 receptor, and  
CC the other site (alpha/gamma site) on nicotinic receptors from  
CC Torpedo californica electric organ.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.  
CC -1- SIMILARITY: Belongs to the conotoxin A-superfamily. Alpha-type  
CC family.  
DR PIR; A01782; NTRNAG.  
DR PDB; 1NOT; X-ray; @=1-14.  
DR PDB; 1QS3; NMR; A=2-13.  
DR PDB; 1XGA; NMR; @=1-14.  
DR PDB; 1XGB; NMR; @=1-14.  
DR PDB; 1XGC; NMR; @=1-14.  
KW 3D-structure; Acetylcholine receptor inhibitor; Amidation;  
KW Direct protein sequencing; Neurotoxin; Postsynaptic neurotoxin; Toxin.  
FT PEPTIDE 1 15 Alpha-conotoxin GIA.  
FT PEPTIDE 1 13 Alpha-conotoxin GI.  
FT DISULFID 2 7  
FT DISULFID 3 13  
FT MOD\_RES 13 13  
FT MOD\_RES 15 15 Cysteine amide (G-14 provides amide  
FT MUTAGEN 9 9 group) (in alpha-conotoxin GI).  
FT STRAND 2 2 Lysine amide (in alpha-conotoxin GIA).  
FT HELIX 5 10 R->A: Reduction in affinity for both  
FT STRAND 12 12 alpha/delta and alpha/gamma sites on  
FT SEQUENCE 15 AA; 1628 MW; 2AE73EE90F8C2B19 CRC64; Torpedo receptors (in GI).  
Query Match 31.6%; Score 6; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 4.9;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 10 CCNPAC 15  
Db 2 CCNPAC 7  
RESULT 14  
Q9R581 PRELIMINARY; PRT; 17 AA.  
AC Q9R581;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE 01-ST-1, NAG-ST, VM-ST=HEAT-stable enterotoxin.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=666;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=93314823; PubMed=8325391; DOI=10.1016/0014-5793(93)81766-S;  
RA Yoshino K., Miyachi M., Takao T., Bag P.K., Huang X., Nair G.B.,  
RA Takeda T., Shimonishi Y.;  
RT "Purification and sequence determination of heat-stable enterotoxin  
RT elaborated by a cholera toxin-producing strain of Vibrio cholerae  
RT O1.";  
RL FEBS Lett. 326:83-86 (1993).  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0009405; P:Pathogenesis; IEA.  
DR InterPro; IPR001489; Enterotoxin HS.  
DR Pfam; PF02048; Enterotoxin HS; 1.  
DR PROSITE; PS00273; ENTEROTOXIN H STABLE; 1.  
DR SEQUENCE 17 AA; 1821 MW; 30FF036D018D601C CRC64;

Query Match 31.6%; Score 6; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 5.4;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 CCNPAC 15  
Db 7 CCNPAC 12

## RESULT 15

Q9R580 PRELIMINARY; PRT; 18 AA.  
AC Q9R580;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE 01-ST-2, VC-H-ST=HEAT-stable enterotoxin.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=666;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=93314823; PubMed=8325391; DOI=10.1016/0014-5793(93)81766-S;  
RA Yoshino K., Miyachi M., Takao T., Bag P.K., Huang X., Nair G.B.,  
RA Takeda T., Shimonishi Y.;  
RT "Purification and sequence determination of heat-stable enterotoxin  
RT elaborated by a cholera toxin-producing strain of Vibrio cholerae  
RT O1";  
RL FEBS Lett. 326:83-86(1993).  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0009405; P:pathogenesis; IEA.  
DR InterPro; IPR001489; Enterotoxin HS.  
DR Pfam; PF02048; Enterotoxin HS; 1.  
DR PROSITE; PS00273; ENTEROTOXIN\_H\_STABLE; 1.  
SQ SEQUENCE 18 AA; 1934 MW; 3080692D018D601C CRC64;

Query Match 31.6%; Score 6; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 5.6;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 CCNPAC 15  
Db 8 CCNPAC 13

Search completed: March 26, 2005, 17:24:50  
Job time : 75.9821 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2005, 16:49:45 ; Search time 80.0357 Seconds  
(without alignments)  
86.982 Million cell updates/sec

Title: US-10-775-481A-3  
Perfect score: 18  
Sequence: 1 NTFYCCCLCCVPACAGCN 18

Scoring table:  Gapop 60.0, Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID      | Description        |
|------------|-------|-------------|--------|------------|--------------------|
| 1          | 18    | 100.0       | 18     | 1 AAP40186 | Aap40186 Sequence  |
| 2          | 18    | 100.0       | 18     | 2 Aar85947 | Aar85947 ST I* rec |
| 3          | 18    | 100.0       | 18     | 2 Aay40509 | Aay40509 ST recept |
| 4          | 18    | 100.0       | 18     | 5 Aau93984 | Aau93984 E. coli S |
| 5          | 18    | 100.0       | 18     | 7 Ade10846 | Ade10846 Chimeric  |
| 6          | 18    | 100.0       | 18     | 7 Adm39751 | Adm39751 Escherich |
| 7          | 18    | 100.0       | 18     | 8 Adg63946 | Adg63946 Recombina |
| 8          | 18    | 100.0       | 18     | 8 Adp73480 | Adp73480 Escherich |
| 9          | 18    | 100.0       | 18     | 8 Adr12773 | Adr12773 E. coli S |
| 10         | 18    | 100.0       | 18     | 8 Aar84343 | Aar84343 Bacterial |
| 11         | 18    | 100.0       | 18     | 8 Adr48331 | Adr48331 E. coli S |
| 12         | 18    | 100.0       | 18     | 8 Aar45823 | Aar45823 Amino aci |
| 13         | 18    | 100.0       | 36     | 1 AAP30263 | Aap30263 Sequence  |
| 14         | 17    | 94.4        | 17     | 2 Aar85966 | Aar85966 ST I* rec |
| 15         | 17    | 94.4        | 17     | 2 Aar85961 | Aar85961 ST I* rec |
| 16         | 17    | 94.4        | 17     | 2 Aay40528 | Aay40528 ST recept |
| 17         | 17    | 94.4        | 17     | 2 Aay40523 | Aay40523 ST recept |
| 18         | 17    | 94.4        | 17     | 8 Adr45837 | Adr45837 Amino aci |
| 19         | 17    | 94.4        | 17     | 8 Aar45842 | Aar45842 Amino aci |
| 20         | 16    | 88.9        | 16     | 2 Aar85962 | Aar85962 ST I* rec |
| 21         | 16    | 88.9        | 16     | 2 Aar85967 | Aar85967 ST I* rec |
| 22         | 16    | 88.9        | 16     | 2 Aay40524 | Aay40524 ST recept |
| 23         | 16    | 88.9        | 16     | 2 Aay40529 | Aay40529 ST recept |
| 24         | 16    | 88.9        | 16     | 8 Adr45838 | Adr45838 Amino aci |
| 25         | 16    | 88.9        | 16     | 8 ADR45843 | ADR45843 Amino aci |

|    |    |      |    |            |                    |
|----|----|------|----|------------|--------------------|
| 26 | 15 | 83.3 | 15 | 2 AAR85968 | Aar85968 ST I* rec |
| 27 | 15 | 83.3 | 15 | 2 Aar85963 | Aar85963 ST I* rec |
| 28 | 15 | 83.3 | 15 | 2 Aay40525 | Aay40525 ST recept |
| 29 | 15 | 83.3 | 15 | 2 Aay40530 | Aay40530 ST recept |
| 30 | 15 | 83.3 | 15 | 8 ADR45844 | ADR45844 Amino aci |
| 31 | 15 | 83.3 | 15 | 8 ADR45839 | ADR45839 Amino aci |
| 32 | 15 | 83.3 | 18 | 5 AAU93985 | Aau93985 E. coli S |
| 33 | 15 | 83.3 | 18 | 7 ADE10847 | Adel10847 Chimeric |
| 34 | 15 | 83.3 | 18 | 7 ADM39752 | Adm39752 Escherich |
| 35 | 15 | 83.3 | 18 | 8 ADG63947 | Adg63947 Recombina |
| 36 | 15 | 83.3 | 18 | 8 ADP73481 | Adp73481 Escherich |
| 37 | 15 | 83.3 | 18 | 8 ADR12774 | Adr12774 E. coli S |
| 38 | 14 | 77.8 | 14 | 2 AAR85964 | Aar85964 ST I* rec |
| 39 | 14 | 77.8 | 14 | 2 AAR85969 | Aar85969 ST I* rec |
| 40 | 14 | 77.8 | 14 | 2 AAY40531 | Aay40531 ST recept |
| 41 | 14 | 77.8 | 14 | 2 AAY40526 | Aay40526 ST recept |
| 42 | 14 | 77.8 | 14 | 5 AAU93983 | Aau93983 E. coli S |
| 43 | 14 | 77.8 | 14 | 7 ADE10845 | Ade10845 Chimeric  |
| 44 | 14 | 77.8 | 14 | 7 ADM39750 | Adm39750 Escherich |
| 45 | 14 | 77.8 | 14 | 8 ADG63945 | Adg63945 Recombina |

#### ALIGNMENTS

RESULT 1  
AAP40186  
ID AAP40186 standard; peptide; 18 AA.  
XX  
AC AAP40186;  
XX  
DT 25-MAR-2003 (revised)  
DT 04-FEB-1992 (first entry)  
XX  
DE Sequence of monomeric synthetic heat-stable Escherichia coli enterotoxin  
DE IB analogue VII.

Vaccine; diarrhoea; diagnosis; antigen.

Escherichia coli.

W08402700-A.

19-JUL-1984.

21-DEC-1983; 83WO-US002008.

03-JAN-1983; 83US-00455265.

12-DEC-1983; 83US-00559469.

12-DEC-1984; 84WO-US002030.  
(SCRI ) SCRIPPS CLINIC & RES FOUND.

Houghten RA;

WPI; 1984-188753/30.

Synthetic polypeptide(s) with antigenicity of E.coli enterotoxin - useful for vaccination against diarrhoea and for diagnostic tests.

Example; Page 53; 177pp; English.

The peptides of the invention can be used in a vaccine useful for conferring protection against diarrhoea. They may also be used in a diagnostic or reagent system for assaying for infections caused by the E.coli strains. (Updated on 25-MAR-2003 to correct PA field.)

Sequence 18 AA;

Query Match 100.0%; Score 18; DB 1; Length 18;

Best Local Similarity 100.0%; Pred. No. 1.3e-12; Indels 0; Gaps 0;  
Matches 18; Conservative 0; Mismatches 0;

```

QY      1 NTFYCCCLCCYPACAGCN 18
Db      1 NTFYCCCLCCYPACAGCN 18

RESULT 2
AAR85947
ID AAR85947 standard; peptide; 18 AA.
XX
AC AAR85947;
XX
DT 19-JAN-1996 (first entry)
XX
DE ST I* receptor ligand portion.
XX
KW ST; receptor; ligand; heat-stable toxin; conjugate; imaging; cancer;
KW colorectal; metastasis.
XX
OS Escherichia coli.
XX
PN W09511694-A1.
XX
PD 04-MAY-1995.
XX
PF 26-OCT-1994; 94WO-US012232.
XX
PR 26-OCT-1993; 93US-00141892.
PR 13-SEP-1994; 94US-00305056.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Waldman SA;
XX
DR WPI; 1995-178646/23.
XX
DT Conjugated cpds. which specifically bind to colorectal cancer cells -
PT comprise heat-stable toxin receptor binding moiety and active moiety
PT which may be a therapeutic agent or a radioactive agent.
XX
PS Claim 3; Page 116; 133pp; English.
XX
CC New conjugated compounds are provided which consist of (1) an ST receptor
CC binding moiety and (2) an active moiety which is a radio- stable agent.
CC 'ST' refers to a heat stable toxin produced by E.coli and other
CC organisms. Especially the radiostable agent is a therapeutic agent (e.g.
CC methotrexate, doxorubicin, mitomycin), imaging agent (e.g. radioactive
CC isotope) or nucleic acid; and the compound is used for the detection,
CC imaging or treatment of colorectal tumours, particularly metastasised
CC tumours. The present sequence is a specific example of an ST receptor
CC binding peptide which can be used in the conjugate
XX
SQ Sequence 18 AA;

Query Match      100.0%; Score 18; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.3e-12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 NTFYCCCLCCYPACAGCN 18
Db      1 NTFYCCCLCCYPACAGCN 18

RESULT 4
AAU93984
ID AAU93984 standard; peptide; 18 AA.
XX
AC AAU93984;
XX
DT 02-JUL-2002 (first entry)
XX
DE E. coli ST B cell epitope #2.
XX
KW Immunogenic; hepatitis B core; HBc;
KW vaccine; B cell epitope; T cell epitope; immunostimulant.
XX
OS Escherichia coli.
XX
PN W0200214478-A2.
XX
PD 21-FEB-2002.
XX
PF 16-AUG-2001; 2001WO-US041759.
XX
PR 16-AUG-2000; 2000US-0225843P.
PR 22-AUG-2000; 2000US-0226867P.
PR 15-AUG-2001; 2001US-00930915.
XX
PA (APOV-) APOVIA INC.
XX
PI Birkett AJ;
XX
DR WPI; 2002-257601/30.

```

PT Novel recombinant hepatitis B nucleocapsid protein, termed as chimeric  
 PT hepatitis B core protein, displays immunogenic epitopes at N-terminus,  
 PT HBC immunogenic loop with linker for conjugated epitope and C-terminus.  
 XX  
 PS Disclosure; Page 38; 289pp; English.

XX  
 CC The invention relates to a recombinant hepatitis B nucleocapsid protein,  
 CC i.e. a chimeric hepatitis B core (HBC) protein (I), displaying one or  
 CC more immunogenic epitopes at the N-terminus, HBC immunogenic loop (L) or  
 CC C-terminus, or having a heterologous linker for a conjugated epitope in  
 CC (L), and containing a Cys residue at, or near, the C-terminus that  
 CC confers enhanced stability to the particles. A vaccine comprising (I) is  
 CC useful for inducing an immune response in an inoculated host animal, by  
 CC inoculating a host animal with the vaccine, and maintaining that  
 CC inoculated animal for a time period sufficient for that animal to develop  
 CC an immune response. The immunogenic particles formed using (I) are  
 CC substantially free of binding to nucleic acids, and are most stable than  
 CC the particle formed from otherwise identical HBC chimera that lacks the C-  
 CC terminal residue or in which a C-terminal Cys is replaced by another  
 CC residue. The chimera particles are most stable on storage in aqueous  
 CC compositions that are particles of similar sequence that lack any C-  
 CC terminal Cys residues. The chimera molecule exhibits the self-assembly not  
 CC exhibiting the nucleic acid binding of those native particles, and  
 CC excellent B cell and T cell immunogenicities. The chimera particles are  
 CC typically prepared in higher yield than similar particles that are free  
 CC of a C-terminal Cys. The particles are often far more immunogenic than  
 CC the similar conjugates that lack a C-terminal Cys. Immunogenicities of  
 CC particles assembled from the chimera molecules are enhanced as compared to  
 CC similar particles assembled from chimera molecules lacking at least one C-  
 CC terminal Cys. AAU93802-AAU93997 represent immunogenic HBC particles amino  
 CC acid sequences and related sequences of the invention

XX  
 SQ Sequence 18 AA;

Query Match 100.0%; Score 18; DB 5; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-12;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NTFYCCELCYCYPACAGCN 18  
 |||||  
 Db 1 NTFYCCELCYCYPACAGCN 18

RESULT 5  
 ADE10846

ID ADE10846 standard; peptide; 18 AA.

XX AC ADE10846;

XX DT 29-JAN-2004 (first entry)

XX DE Chimeric hepatitis B virus related B-cell epitope seqid 80.

XX KW hepatotropic; virucide; antiinflammatory; chronic hepatitis; vaccine;  
 KW recombinant hepatitis B core chimeric protein; HBC chimeric protein;  
 KW hepatitis B infection; T-cell stimulator; B-cell epitope.

XX OS Escherichia coli.

XX PN US2003198645-A1.

XX PD 23-OCT-2003.

XX PF 21-FEB-2003; 2003US-00372076.

XX PR 21-FEB-2002; 2002US-00080299.

XX PR 21-FEB-2002; 2002US-00082014.

XX (PAGE/) PAGE M.

XX PA (FRIE/) FRIEDE M.

XX PI Page M, Friede M;

XX

DR WPI; 2003-852775/79.

XX  
 PT Treating chronic hepatitis B infection by administering a T cell-  
 PT stimulating vaccine containing immunogenic particles having recombinant  
 PT carboxy-terminal truncated hepatitis B core (HBC) chimeric protein  
 PT molecules.

XX PS Disclosure; SEQ ID NO 80; 111pp; English.

XX  
 CC The invention describes a method of treating chronic hepatitis comprising  
 CC administering to a patient a T cell-stimulating amount of a vaccine  
 CC comprising immunogenic particles dissolved or dispersed in a diluent,  
 CC where the immunogenic particles consists of recombinant hepatitis B core  
 CC (HBC) chimeric protein molecules, and maintaining the patient to induce T  
 CC cells activated against HBC. The methods and compositions of the present  
 CC invention are useful for treating chronic hepatitis B infection. This is  
 CC the amino acid sequence of a chimeric hepatitis B virus related B-cell  
 CC epitope useful for expression within the HBV chimera at the N-terminus,  
 CC within the immunogenic loop and/or at the C-terminus.

XX SQ Sequence 18 AA;

Query Match 100.0%; Score 18; DB 7; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-12;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NTFYCCELCYCYPACAGCN 18

|||||

Db 1 NTFYCCELCYCYPACAGCN 18

RESULT 6

ADM39751

ID ADM39751 standard; peptide; 18 AA.

XX AC ADM39751;

XX DT 03-JUN-2004 (first entry)

XX DE Escherichia coli B-cell peptide epitope expressed by HBC chimera Seq 83.

XX KW immunogenic; avian hepatitis B virus; nucleocapsid;

XX KW self assembled particle; immunogen; inoculum; vaccine; immunostimulant;

XX KW antibacterial; virucidal; B-cell epitope.

XX OS Escherichia coli.

XX PN WO2003072722-A2.

XX PD 04-SEP-2003.

XX PF 21-FEB-2003; 2003WO-US005315.

XX PR 21-FEB-2002; 2002US-03591299.

XX PA (APOV-) APOVIA INC.

XX PI Birkett AJ, Peck B;

XX PD WPI; 2003-679948/64.

XX  
 PT New recombinant chimera avian hepatitis B core protein molecule, useful as  
 PT an immunogen for inducing a B cell or T cell response to produce  
 PT antibodies, or as a vaccine against pathogens.

XX PS Disclosure; SEQ ID NO 83; 278pp; English.

XX  
 CC This invention relates to novel recombinant immunogenic chimeric avian  
 CC hepatitis B core (AHBC) nucleocapsid proteins. Specifically, it refers to  
 CC an AHBC protein that has been engineered to display an immunogenic B cell  
 CC or T cell epitope, exhibit enhanced stability and an absence of nucleic  
 CC acid binding as a self assembled particle. The present invention  
 CC describes the chimeric AHBC protein as truncated at the C-terminus and

CC containing introduced cysteine residues that confers an enhanced  
 CC stability in aqueous solution, an increased yield and more immunogenicity  
 CC than similar conjugates that lack N- or C-terminal cysteines.  
 CC Furthermore, a reduction in the number of positively charged residues  
 CC (lysine and arginine) towards the C-terminus prepares self-assembled  
 CC particles that are substantially free of nucleic acid binding. As such,  
 CC these chimeric particles can be used as immunogens of an inoculum that  
 CC induce a B cell or T cell response in an animal to produce antibodies. It  
 CC can also be useful for developing a vaccine to protect against the  
 CC pathogen from which the heterologous epitope or the haptens is derived.  
 CC Accordingly, these compositions exhibit immunostimulant, antibacterial  
 CC and virucidal activities. This peptide sequence is an exemplary B-cell  
 CC epitope peptide immunogen useful for both linkage to the linker residue  
 CC after expression of a contemplated chimera and for expression within an  
 CC HBC chimera of the invention.

XX Sequence 18 AA;  
 SQ Query Match 100.0%; Score 18; DB 7; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-12;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTFYCCCLCCYPACACGN 18  
 |||||  
 Db 1 NTFYCCCLCCYPACACGN 18

## RESULT 7

ADG63946  
 ID ADG63946 standard; peptide; 18 AA.

XX AC ADG63946;

XX DT 11-MAR-2004 (first entry)

XX DE Recombinant chimera hepatitis B core protein immunogenic epitope #71.  
 XX KW Recombinant chimera hepatitis B core protein; HBC; immunogenic epitope;  
 XX KM HBC immunodominant loop; immune response.

XX OS Escherichia coli.

XX PN US2003185858-A1.

XX PD 02-OCT-2003.

XX PF 21-FEB-2002; 2002US-00082014.

XX PR 15-AUG-2001; 2001US-00930915.

XX PA (BIRK/) BIRKETT A J.

XX PI Birkett AJ;

XX DR WPI; 2004-031988/03.

XX PT Recombinant chimera hepatitis B core protein molecule useful for preparing  
 PT vaccine or inoculum includes peptide-bonded heterologous immunogenic  
 PT epitope at N-terminus in the hepatitis B core immunodominant loop or C-  
 PT terminus of the chimera.

XX PS Disclosure; SEQ ID NO 79; 110pp; English.

XX CC The invention relates to a recombinant chimera hepatitis B core (HBC)  
 CC protein molecule that includes a peptide-bonded heterologous immunogenic  
 CC epitope at the N-terminus in the HBC immunodominant loop or the C-  
 CC terminus of the chimera, or a heterologous linker residue to a conjugated  
 CC epitope present in the loop. The invention also relates to an immunogenic  
 CC particle comprising the recombinant hepatitis B core chimera protein  
 CC molecules, a vaccine comprising the immunogenic particles dissolved or  
 CC dispersed in a diluent, a nucleic acid that encodes a recombinant HBC  
 CC protein molecule or its variant, analogue, or complement and a method for  
 CC inducing an immune response in an inoculated host animal comprising

CC inoculating a host animal with a vaccine and maintaining the inoculated  
 CC animal for a period of time sufficient to enable development of an immune  
 CC response. The recombinant chimera hepatitis B core protein molecule is  
 CC used in an immunogenic particle for preparing a vaccine useful for  
 CC inducing an immune response in an inoculated host animal. This sequence  
 CC represents an HBC protein immunogenic B cell epitope of the invention.

XX Sequence 18 AA;

XX Query Match 100.0%; Score 18; DB 8; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-12;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTFYCCCLCCYPACACGN 18  
 |||||  
 Db 1 NTFYCCCLCCYPACACGN 18

## RESULT 8

ADP73480  
 ID ADP73480 standard; peptide; 18 AA.

XX AC ADP73480;

XX DT 09-SEP-2004 (first entry)

XX DE Escherichia coli B cell epitope of gene ST, SEQ ID 93.

XX KW transgenic animal; Hepatitis B virus nucleocapsid core protein; HBC;  
 XX KM enhanced stability; hepatotropic; virucide; immunology;  
 XX KW protein engineering; immunogen; vaccine; Hepatitis B infection.

XX OS Escherichia coli.

XX PN WO2004053091-A2.

XX PD 24-JUN-2004.

XX PF 10-DEC-2003; 2003WO-US039164.

XX PR 10-DEC-2002; 2002US-0432123P.

XX PA (APOV-) APOVIA INC.

XX PI Lyons K, Birkett AJ, Haron JA;

XX DR WPI; 2004-468859/44.

XX PT New recombinant chimera hepatitis B core (HBC) protein molecules useful in  
 PT the fields of immunology and protein engineering, in particular as an  
 PT immunogen in a vaccine for Hepatitis B infections.

XX PS Disclosure; SEQ ID NO 93; 338pp; English.

XX CC The invention relates to a novel recombinant chimera hepatitis B virus  
 CC nucleocapsid (core) protein (HBC), up to 600 or 380 amino acid residues  
 CC in length. The chimera protein is engineered for both enhanced stability  
 CC of self-assembled particles and the substantial absence of nucleic acid  
 CC binding by the particles. The invention further comprises: a recombinant  
 CC HBC protein chimera molecule that has a length of 135-365 amino acid  
 CC residues and contains four peptide-linked amino acid residue sequence  
 CC domains from the N-terminus that are denominated Domains I, II, III and  
 CC IV. The invention also provides nucleic acids, polypeptides, host cells,  
 CC vectors and transgenic animals used in the methods of the invention. The  
 CC chimeric compositions of the invention have hepatotropic and virucide  
 CC activities. The methods and compositions of the present invention are  
 CC useful in the fields of immunology and protein engineering, in particular  
 CC for using a chimera hepatitis B virus nucleocapsid protein as an  
 CC immunogen in a vaccine for Hepatitis B infections. This sequence  
 CC represents a hepatitis B virus nucleocapsid (core) protein related  
 CC polypeptide of the invention.

XX Sequence 18 AA;

Query Match 100.0%; Score 18; DB 8; Length 18;  
Best Local Similarity 100.0%; Pred. No. 1.3e-12;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTFYCCCLCCYPACAGCN 18  
| | | | | | | | | | | | | | | | | |  
DB 1 NTFYCCCLCCYPACAGCN 18  
| | | | | | | | | | | | | | | | | |

RESULT 9  
ID ADR12773 standard; peptide; 18 AA.  
XX ADR12773;  
AC 04-NOV-2004 (first entry)  
DT E. coli ST B cell epitope #2.  
DE  
XX HBV; chronic hepatitis; Hbc; nucleocapsid core protein; vaccine;  
KW immunogen; immunogenic epitope; T cell; B cell; CD8+ cell; CD4+ cell;  
KW cytotoxic T lymphocyte; toll-like receptor 4; toll-like receptor 9;  
KW TLR-4; TLR-9; epitope.  
XX  
OS Escherichia coli.  
XX  
PN US2004156863-A1.  
XX  
PD 12-AUG-2004.  
XX  
PF 01-OCT-2003; 2003US-00677074.  
XX  
PR 21-FEB-2002; 2002US-00080299.  
XX  
PR 21-FEB-2002; 2002US-00082014.  
XX  
PR 21-FEB-2003; 2003US-00372076.  
XX  
XX (PAGE// PAGE M.  
PA (FRIE// FRIEDE M.  
PA (SCHM// SCHMIDT A E.  
PA (STOB// STOB D.  
XX  
XX Page M, Friede M, Schmidt AE, Stober D;  
XX WPI; 2004-603322/58.  
XX  
XX Treating chronic hepatitis, by administering vaccine comprising  
PT immunogenic particles having recombinant hepatitis B core chimeric  
PT protein molecules, that stimulates T cell, to patient chronically  
PT infected with hepatitis B virus.  
XX  
XX Disclosure; SEQ ID NO 80; 117pp; English.

The invention relates to treating chronic hepatitis, by administering a vaccine comprising immunogenic particles having recombinant hepatitis B core (Hbc) chimeric protein molecules (where truncated Hbc molecules are linked N-terminally or C-terminally to an immunogenic epitope), that stimulate T cell production, to a patient chronically infected with hepatitis B virus, and maintaining patient for time sufficient to induce T cells activated against Hbc. The chimeric proteins are still capable self-assembling into particles upon expression in a host cell and are substantially free of binding to nucleic acids, and the particles display enhanced stability. Also included is enhancing (M2) the production of one or more of gamma-producing CD8+, CD4+ T cells and cytotoxic T lymphocytes against hepatitis B virus, involving administering to a patient chronically infected with hepatitis B virus, a T cell-stimulating amount of a vaccine comprising immunogenic particles dissolved or dispersed in a diluent containing one or both of an agonist of toll-like receptor 4 and receptor 9 (TLR-4 and TLR-9), the immunogenic particles comprising Hbc chimeric protein molecules and maintaining the patient for a sufficient time to induce T cells activated against Hbc. The immunogenic epitopes may be B cell or T cell epitopes. The chimeric vaccine is useful for treating a patient chronically infected with hepatitis B virus. The

CC present sequence is a B cell epitope suitable for inclusion in the  
CC chimeric protein of the invention.  
XX  
SQ Sequence 18 AA;

Query Match 100.0%; Score 18; DB 8; Length 18;  
Best Local Similarity 100.0%; Pred. No. 1.3e-12;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTFYCCCLCCYPACAGCN 18  
| | | | | | | | | | | | | | | | | |  
DB 1 NTFYCCCLCCYPACAGCN 18  
| | | | | | | | | | | | | | | | | |

RESULT 10  
ID ADR48343 standard; peptide; 18 AA.  
XX ADR48343;  
AC 04-NOV-2004 (first entry)  
DT Bacterial ST peptide #4.  
DE  
XX Gastrointestinal; antiinflammatory; laxative; cardiant; antiulcer;  
KW anorectic; cardiovascular; cytostatic; analgesic; CNS; respiratory;  
KW neuroprotective; vasotropic; auditory; antiemetic; antiasthmatic;  
KW nephrotropic; hepatotropic; virucide; immunosuppressive; antiallergic;  
KW antidiabetic; ophthalmological; tranquiliser; hypnotic; nootropic;  
KW guanylate cyclase C; GC-C; receptor; gastrointestinal disorder;  
KW irritable bowel syndrome; constipation; gastroesophageal reflux disease;  
KW heartburn; dyspepsia; gastroparesis; Crohn's disease; ulcerative colitis;  
KW inflammatory bowel disease; obesity; heart failure; cystic fibrosis;  
KW cancer; respiratory disorder; neurological disorder; carbonate imbalance;  
KW erectile dysfunction; inner ear disorder; slow digestion; nausea;  
KW vomiting; bloating; asthma; nephritis; hepatitis; pancreatitis; allergy;  
KW retinopathy; nephropathy; headache; anxiety; sleep disorder.  
XX  
OS Escherichia coli.  
XX  
PN WO2004069165-A2.  
XX  
PD 19-AUG-2004.  
XX  
XX 28-JAN-2004; 2004WO-US002390.  
XX  
XX 28-JAN-2003; 2003US-0443098P.  
PR 15-MAY-2003; 2003US-0471288P.  
PR 12-NOV-2003; 2003US-0519460P.  
XX  
XX (MICR-) MICROBIA INC.  
XX  
XX Currie MG, Mahajan-Miklos S;  
XX WPI; 2004-604332/58.  
XX  
XX Novel purified peptide capable of activating the guanylate cyclase C  
PT receptor, useful for treating obesity, congestive heart failure and  
PT benign prostatic hyperplasia.  
XX  
XX Disclosure; Page 29; 93pp; English.

The invention relates to a purified peptide (P1) capable of activating the guanylate cyclase C (GC-C) receptor. Further disclosed is a pharmaceutical composition comprising the peptide of the invention. The composition of the invention is useful for treating a gastrointestinal disorder in a patient, which involves administering P1, where the gastrointestinal disorder is gastrointestinal motility disorder, irritable bowel syndrome, chronic constipation, a functional gastrointestinal disorder, gastroesophageal reflux disease, functional heartburn, dyspepsia, functional dyspepsia, nonulcer dyspepsia, gastroparesis, chronic intestinal pseudo-obstruction, colonic pseudo-obstruction, Crohn's disease, ulcerative colitis or inflammatory bowel

CC disease. The peptide of the invention is also useful for treating  
 CC obesity, congestive heart failure, cystic fibrosis or a patient suffering  
 CC from constipation. The P1/GC-C receptor agonist is useful for treating  
 CC cancer, respiratory disorder, neurological disorder, disorder associated  
 CC with carbonate imbalance, erectile dysfunction, insulin-related disorder  
 CC or inner ear disorder. P1 is useful in treating slow digestion or slow  
 CC stomach emptying. P1 is useful in relieving symptoms of gastroparesis  
 CC such as nausea, vomiting, bloating, and delayed gastric emptying. P1 is  
 CC useful for treating or preventing asthma, nephritis, hepatitis,  
 CC pancreatitis, allergies, etc. P1 is useful for treating or preventing  
 CC type II diabetes mellitus, hyperglycaemia, respiratory disorders  
 CC including inhalation. P1 is useful in treating or preventing retinopathy,  
 CC nephropathy and edema formation. P1 is useful for treating or preventing  
 CC headache, anxiety, sleep disorders and memory loss. P1 is useful as a  
 CC marker to identify, detect, stage, or diagnosis diseases and conditions  
 CC of the small intestine, including Crohn's disease, colitis, inflammatory  
 CC bowel disease, tumours, etc. P1 can be conjugated to diagnostic or  
 CC therapeutic molecule to target cells bearing GC-C receptor, e.g., cystic  
 CC fibrosis lesions and specific cells lining the intestinal tract, thus  
 CC useful in targeting radioactive moieties or therapeutic moieties to the  
 CC intestine to aid in imaging and diagnosing or treating  
 CC colorectal/metastasis or local colorectal cancer. The current sequence  
 CC represents a bacterial ST peptide which is an activator of the GC-C  
 CC receptor. ST peptides are considered super agonists of GC-C and are very  
 CC resistant to proteolytic degradation.

SQ Sequence 18 AA;

Query Match 100.0%; Score 18; DB 8; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-12;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTFYCCCLCCYPACAGCN 18  
 DB 1 NTFYCCCLCCYPACAGCN 18

RESULT 11

ADRA48331

XX ADR48331 standard; peptide; 18 AA.

AC ADR48331;

DT 04-NOV-2004 (first entry)

XX E. coli ST I\* peptide.

KW Gastrointestinal; antiinflammatory; laxative; cardiant; antiulcer;  
 KW anorectic; cardiovascular; cytostatic; analgesic; CNS; respiratory;  
 KW neuroprotective; vasotropic; auditory; antiemetic; antiasthmatic;  
 KW nephrotropic; hepatotropic; virucide; immunosuppressive; antiallergic;  
 KW antidiabetic; ophthalmological; tranquiliser; hypnotic; nootropic;  
 KW guanylate cyclase C; GC-C; receptor; gastrointestinal disorder;  
 KW irritable bowel syndrome; constipation; gastroesophageal reflux disease;  
 KW heartburn; dyspepsia; gastroparesis; Crohn's disease; ulcerative colitis;  
 KW inflammatory bowel disease; obesity; heart failure; cystic fibrosis;  
 KW cancer; respiratory disorder; neurological disorder; carbonate imbalance;  
 KW erectile dysfunction; inner ear disorder; slow digestion; nausea;  
 KW vomiting; bloating; asthma; nephritis; hepatitis; pancreatitis; allergy;  
 KW retinopathy; nephropathy; headache; anxiety; sleep disorder.

XX Escherichia coli.

XX WO2004069165-A2.

XX 19-AUG-2004.

XX 28-JAN-2004; 2004WO-US002390.

XX 28-JAN-2003; 2003US-0443098P.

PR 15-MAY-2003; 2003US-0471288P.

PR 12-NOV-2003; 2003US-0519460P.

XX

PA (MICR-) MICROBIA INC.

XX Currie MG, Mahajan-Miklos S;

XX WPI; 2004-604332/58.

DR Novel purified peptide capable of activating the guanylate cyclase C  
 XX receptor, useful for treating obesity, congestive heart failure and  
 PT benign prostatic hyperplasia.

XX Disclosure; Page 28; 93pp; English.

PS The invention relates to a purified peptide (P1) capable of activating  
 XX the guanylate cyclase C (GC-C) receptor. Further disclosed is a  
 CC pharmaceutical composition comprising the peptide of the invention. The  
 CC composition of the invention is useful for treating a gastrointestinal  
 CC disorder in a patient, which involves administering P1, where the  
 CC gastrointestinal disorder is gastrointestinal motility disorder,  
 CC irritable bowel syndrome, chronic constipation, a functional  
 CC gastrointestinal disorder, gastroesophageal reflux disease, functional  
 CC heartburn, dyspepsia, functional dyspepsia, nonulcer dyspepsia,  
 CC gastroparesis, chronic intestinal pseudo-obstruction, colonic pseudo-  
 CC obstruction, Crohn's disease, ulcerative colitis or inflammatory bowel  
 CC disease. The peptide of the invention is also useful for treating  
 CC obesity, congestive heart failure, cystic fibrosis or a patient suffering  
 CC from constipation. The P1/GC-C receptor agonist is useful for treating  
 CC cancer, respiratory disorder, neurological disorder, disorder associated  
 CC with carbonate imbalance, erectile dysfunction, insulin-related disorder  
 CC or inner ear disorder. P1 is useful in treating slow digestion or slow  
 CC stomach emptying. P1 is useful in relieving symptoms of gastroparesis  
 CC such as nausea, vomiting, bloating, and delayed gastric emptying. P1 is  
 CC useful for treating or preventing asthma, nephritis, hepatitis,  
 CC pancreatitis, allergies, etc. P1 is useful for treating or preventing  
 CC type II diabetes mellitus, hyperglycaemia, respiratory disorders  
 CC including inhalation. P1 is useful in treating or preventing retinopathy,  
 CC nephropathy and edema formation. P1 is useful for treating or preventing  
 CC headache, anxiety, sleep disorders and memory loss. P1 is useful as a  
 CC marker to identify, detect, stage, or diagnosis diseases and conditions  
 CC of the small intestine, including Crohn's disease, colitis, inflammatory  
 CC bowel disease, tumours, etc. P1 can be conjugated to diagnostic or  
 CC therapeutic molecule to target cells bearing GC-C receptor, e.g., cystic  
 CC fibrosis lesions and specific cells lining the intestinal tract, thus  
 CC useful in targeting radioactive moieties or therapeutic moieties to the  
 CC intestine to aid in imaging and diagnosing or treating  
 CC colorectal/metastasis or local colorectal cancer. The current sequence  
 CC represents a bacterial ST peptide which is an activator of the GC-C  
 CC receptor. ST peptides are considered super agonists of GC-C and are very  
 CC resistant to proteolytic degradation.

SQ Sequence 18 AA;

Query Match 100.0%; Score 18; DB 8; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-12;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTFYCCCLCCYPACAGCN 18

DB 1 NTFYCCCLCCYPACAGCN 18

RESULT 12

ADRA45823

ID ADR45823 standard; peptide; 18 AA.

XX ADR45823;

XX 18-NOV-2004 (first entry)

XX Amino acid sequence of peptide ST I\* inhibiting ST activity.

XX heat stable toxin; ST; ST receptor; colorectal cancer; gastric cancer;  
 KW oesophageal cancer; colorectal cell; gastric cell; oesophageal cell;  
 KW ST Ia.



PD 09-NOV-1983.

XX New conjugated compounds are provided which consist of (1) an ST receptor  
 CC binding moiety and (2) an active moiety which is a radio- stable agent.  
 CC 'ST' refers to a heat stable toxin produced by E.coli and other  
 CC organisms. Especially the radiostable agent is a therapeutic agent (e.g.  
 CC methotrexate, doxorubicin, mitomycin), imaging agent (e.g. radioactive  
 CC isotope) or nucleic acid; and the compound is used for the detection,  
 CC imaging or treatment of colorectal tumours, particularly metastasised  
 CC tumours. The present sequence is a specific example of an ST receptor  
 CC binding peptide which can be used in the conjugate  
 XX

SQ Sequence 17 AA;

Query Match 94.4%; Score 17; DB 2; Length 17;

Best Local Similarity 100.0%; Pred. No. 1.5e-11;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TFCYCCCLCCYPACAGCN 18

Db 1 TFCYCCCLCCYPACAGCN 17

# RESULT 15

AAR85961  
 ID AAR85961 standard; peptide; 17 AA.

XX AC AAR85961;

XX DT 19-JAN-1996 (first entry)

XX DE ST I\* receptor ligand portion.

XX ST; receptor; ligand; heat-stable toxin; conjugate; imaging; cancer;  
 KW colorectal; metastasis.

XX OS Escherichia coli.

XX PN WO9511694-A1.

XX PD 04-MAY-1995.

XX PF 26-OCT-1994; 94WO-US012232.

XX PR 26-OCT-1993; 93US-00141892.

XX PR 13-SEP-1994; 94US-00305056.

XX PA (UYJE-) UNIV JEFFERSON THOMAS.

XX PI Waldman SA;

XX WPI; 1995-178646/23.

XX Conjugated cpds. which specifically bind to colorectal cancer cells -  
 PT comprise heat-stable toxin receptor binding moiety and active moiety  
 PT which may be a therapeutic agent or a radioactive agent.

XX Claim 3; Page 119; 133pp; English.

XX New conjugated compounds are provided which consist of (1) an ST receptor  
 CC binding moiety and (2) an active moiety which is a radio- stable agent.  
 CC 'ST' refers to a heat stable toxin produced by E.coli and other  
 CC organisms. Especially the radiostable agent is a therapeutic agent (e.g.  
 CC methotrexate, doxorubicin, mitomycin), imaging agent (e.g. radioactive  
 CC isotope) or nucleic acid; and the compound is used for the detection,  
 CC imaging or treatment of colorectal tumours, particularly metastasised  
 CC tumours. The present sequence is a specific example of an ST receptor  
 CC binding peptide which can be used in the conjugate  
 XX

SQ Sequence 17 AA;

Query Match 94.4%; Score 17; DB 2; Length 17;

Best Local Similarity 100.0%; Pred. No. 1.5e-11;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTFYCCCLCCYPACAGC 17

Db 1 NTFYCCCLCCYPACAGC 17


Search completed: March 26, 2005, 17:20:57

Job time : 80.0357 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2005, 17:12:01 ; Search time 24.4286 Seconds  
(without alignments)  
55.005 Million cell updates/sec

Title: US-10-775-481A-3  
 Perfect score: 18  
 Sequence: 1 NTFVCELCCYPACAGCN 18  
 Scoring table: 

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 513545

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Listing first 45 summaries

```
Database :
1: /cgn2_6/pdata/1/iaa/5A_COMB pep.*
2: /cgn2_6/pdata/1/iaa/5B_COMB pep.*
3: /cgn2_6/pdata/1/iaa/6A_COMB pep.*
4: /cgn2_6/pdata/1/iaa/6B_COMB pep.*
5: /cgn2_6/pdata/1/iaa/PTTUS_COMB pep.*
6: /cgn2_6/pdata/1/iaa/backfiles1 pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Query |       |        | DB | ID                | Description        |
|------------|-------|-------|--------|----|-------------------|--------------------|
|            | Score | Match | Length |    |                   |                    |
| 1          | 18    | 100.0 | 18     | 1  | US-08-141-892A-3  | Sequence 3, Appli  |
| 2          | 18    | 100.0 | 18     | 2  | US-08-583-447A-3  | Sequence 3, Appli  |
| 3          | 18    | 100.0 | 18     | 2  | US-08-467-920-3   | Sequence 3, Appli  |
| 4          | 18    | 100.0 | 18     | 3  | US-08-635-930-3   | Sequence 3, Appli  |
| 5          | 18    | 100.0 | 18     | 3  | US-09-193-997-3   | Sequence 3, Appli  |
| 6          | 18    | 100.0 | 18     | 3  | US-09-138-237A-3  | Sequence 3, Appli  |
| 7          | 17    | 94.4  | 17     | 1  | US-08-141-892A-18 | Sequence 18, Appli |
| 8          | 17    | 94.4  | 17     | 1  | US-08-141-892A-23 | Sequence 23, Appli |
| 9          | 17    | 94.4  | 17     | 2  | US-08-583-447A-23 | Sequence 18, Appli |
| 10         | 17    | 94.4  | 17     | 2  | US-08-583-447A-23 | Sequence 23, Appli |
| 11         | 17    | 94.4  | 17     | 2  | US-08-467-920-18  | Sequence 18, Appli |
| 12         | 17    | 94.4  | 17     | 2  | US-08-467-920-23  | Sequence 23, Appli |
| 13         | 17    | 94.4  | 17     | 3  | US-08-635-930-18  | Sequence 18, Appli |
| 14         | 17    | 94.4  | 17     | 3  | US-08-635-930-23  | Sequence 23, Appli |
| 15         | 17    | 94.4  | 17     | 3  | US-09-193-997-18  | Sequence 18, Appli |
| 16         | 17    | 94.4  | 17     | 3  | US-09-193-997-23  | Sequence 23, Appli |
| 17         | 17    | 94.4  | 17     | 3  | US-09-138-237A-18 | Sequence 18, Appli |
| 18         | 17    | 94.4  | 17     | 3  | US-09-138-237A-23 | Sequence 23, Appli |
| 19         | 16    | 88.9  | 16     | 1  | US-08-141-892A-19 | Sequence 19, Appli |
| 20         | 16    | 88.9  | 16     | 1  | US-08-141-892A-24 | Sequence 24, Appli |
| 21         | 16    | 88.9  | 16     | 2  | US-08-583-447A-19 | Sequence 19, Appli |
| 22         | 16    | 88.9  | 16     | 2  | US-08-583-447A-24 | Sequence 24, Appli |
| 23         | 16    | 88.9  | 16     | 2  | US-08-467-920-19  | Sequence 19, Appli |
| 24         | 16    | 88.9  | 16     | 2  | US-08-467-920-24  | Sequence 24, Appli |
| 25         | 16    | 88.9  | 16     | 3  | US-08-635-930-19  | Sequence 19, Appli |
| 26         | 16    | 88.9  | 16     | 3  | US-08-635-930-24  | Sequence 24, Appli |
| 27         | 16    | 88.9  | 16     | 3  | US-09-193-997-19  | Sequence 19, Appli |

## ALIGNMENTS

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

|                           |         |                    |       |            |
|---------------------------|---------|--------------------|-------|------------|
| Query Match               | 100.0%; | Score 18;          | DB 1; | Length 18; |
| Best Local Similarity     | 100.0%; | Pred. No. 2.8e-13; |       |            |
| Matches -18: Conservative | 0;      | Mismatches         | 0;    | Indels     |

Qy 1 NTFYCC~~ELCC~~YPACAGCN 18  
|||  
Db 1 NTFYCC~~ELCC~~YPACAGCN 18

```
RESULT 2
US-08-583-447A-3
; Sequence 3, Application US/08583447A
; Patent No. 5879656
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; TITLE OF INVENTION: ST Receptor Binding Compounds and
; TITLE OF INVENTION: Methods of Using the Same
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5879656ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,447A
; FILING DATE: 05-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/141,892
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1702
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-583-447A-3
Query Match 100.0%; Score 18; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 2.8e-13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTFYCCELCCYPACAGCN 18
Db 1 NTFYCCELCCYPACAGCN 18

RESULT 3
US-08-467-920-3
; Sequence 3, Application US/08467920
; Patent No. 5962220
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; TITLE OF INVENTION: Compositions That Specifically
; TITLE OF INVENTION: Bind To Colorectal Cancer Cells
; TITLE OF INVENTION: And Methods Of Using The Same
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103

Qy 1 NTFYCCELCCYPACAGCN 18
Db 1 NTFYCCELCCYPACAGCN 18

RESULT 3
US-08-635-930-3
; Sequence 3, Application US/08635930
; Patent No. 6060037
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; TITLE OF INVENTION: Compositions That Specifically Bind To
; TITLE OF INVENTION: Colorectal Cancer Cells And Methods Of Using
; TITLE OF INVENTION: The Same
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6060037ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: WINDOWS 3.1
; SOFTWARE: WordPerfect 6.0/6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,930
; FILING DATE: 26-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/141,892
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/305,056
; FILING DATE: 13-SEP-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,920
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/141,892
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1589
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-467-920-3
Query Match 100.0%; Score 18; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 2.8e-13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTFYCCELCCYPACAGCN 18
Db 1 NTFYCCELCCYPACAGCN 18

RESULT 4
US-08-635-930-3
; Sequence 3, Application US/08635930
; Patent No. 6060037
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; TITLE OF INVENTION: Compositions That Specifically Bind To
; TITLE OF INVENTION: Colorectal Cancer Cells And Methods Of Using
; TITLE OF INVENTION: The Same
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6060037ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: WINDOWS 3.1
; SOFTWARE: WordPerfect 6.0/6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,930
; FILING DATE: 26-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/141,892
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/305,056
; FILING DATE: 13-SEP-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
```

REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-1360  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-635-930-3

Query Match 100.0%; Score 18; DB 3; Length 18;  
Best Local Similarity 100.0%; Pred. No. 2.8e-13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTFYCCELCCYPACGCN 18  
Db 1 NTFYCCELCCYPACGCN 18

RESULT 5  
US-09-193-997-3  
Sequence 3, Application US/09193997  
Patent No. 6087109  
GENERAL INFORMATION:  
APPLICANT: Waldman, Scott A.  
TITLE OF INVENTION: Compositions That Specifically Bind To Colorectal Cancer Cells  
TITLE OF INVENTION: Bind To Colorectal Cancer Cells  
TITLE OF INVENTION: And Methods Of Using The Same  
NUMBER OF SEQUENCES: 54  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
ADDRESS: No. 6087109ris  
STREET: One Liberty Place, 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/193,997  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/467,920  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Deluca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-1589  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-193-997-3

Query Match 100.0%; Score 18; DB 3; Length 18;  
Best Local Similarity 100.0%; Pred. No. 2.8e-13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTFYCCELCCYPACGCN 18

Db 1 NTFYCCELCCYPACGCN 18  
RESULT 6  
US-09-138-237A-3  
Sequence 3, Application US/09138237A  
Patent No. 6268159  
GENERAL INFORMATION:  
APPLICANT: Waldman, Scott A.  
TITLE OF INVENTION: ST Receptor Binding Compounds and Methods  
TITLE OF INVENTION: Of Using the Same  
NUMBER OF SEQUENCES: 54  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 6268159ris  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch disk, 720 Kb  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/138,237A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/141,892  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Deluca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-0903  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-138-237A-3

Query Match 100.0%; Score 18; DB 3; Length 18;  
Best Local Similarity 100.0%; Pred. No. 2.8e-13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTFYCCELCCYPACGCN 18  
Db 1 NTFYCCELCCYPACGCN 18

RESULT 7  
US-08-141-892A-18  
Sequence 18, Application US/08141892A  
Patent No. 5518888  
GENERAL INFORMATION:  
APPLICANT: Waldman, Scott A.  
TITLE OF INVENTION: ST Receptor Binding Compounds and Methods  
TITLE OF INVENTION: Of Using the Same  
NUMBER OF SEQUENCES: 54  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 5518888ris  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19103  
COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch disk, 720 Kb  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/141,892A  
; FILING DATE: 26-OCT-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: TJU-0903  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-141-892A-18

Query Match 94.4%; Score 17; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3.3e-12;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTFYCCELCCYPACGC 17  
Db 1 NTFYCCELCCYPACGC 17

RESULT 8  
US-08-141-892A-23  
; Sequence 23, Application US/08141892A  
; Patent No. 5518888  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: ST Receptor Binding Compounds and Methods  
; TITLE OF INVENTION: of Using the Same  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 5518888-18  
; STREET: One Liberty Place - 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: U.S.A.  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch disk, 720 Kb  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/141,892A  
; FILING DATE: 26-OCT-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: TJU-0903  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids

; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-141-892A-23

Query Match 94.4%; Score 17; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3.3e-12;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TFYCCELCCYPACGCN 18  
Db 1 TFYCCELCCYPACGCN 17

RESULT 9  
US-08-583-447A-18  
; Sequence 18, Application US/08583447A  
; Patent No. 5879656  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: ST Receptor Binding Compounds and  
; TITLE OF INVENTION: Methods of Using the Same  
; NUMBER OF SEQUENCES: 56  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5879656-18  
; STREET: One Liberty Place, 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: WordPerfect 6.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/583,447A  
; FILING DATE: 05-JAN-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/141,892  
; FILING DATE: 26-OCT-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: TJU-1702  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-583-447A-18

Query Match 94.4%; Score 17; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3.3e-12;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTFYCCELCCYPACGC 17  
Db 1 NTFYCCELCCYPACGC 17

RESULT 10  
US-08-583-447A-23  
; Sequence 23, Application US/08583447A  
; Patent No. 5879656  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.

```

; TITLE OF INVENTION: ST Receptor Binding Compounds and
; TITLE OF INVENTION: Methods of Using the Same
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5879656ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/593,447A
; FILING DATE: 05-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/141,892
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1702
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-583-447A-23

Query Match 94.4%; Score 17; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 3.3e-12;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TFYCCELCCYPACACGN 18
Db 1 TFYCCELCCYPACACGN 17

RESULT 11
US-08-467-920-18
; Sequence 18, Application US/08467920
; Patent No. 5962220
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; TITLE OF INVENTION: Compositions That Specifically
; TITLE OF INVENTION: Bind To Colorectal Cancer Cells
; TITLE OF INVENTION: And Methods Of Using The Same
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: No. 5962220ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,920
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/141,892
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1589
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FILING DATE:

```

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/141,892
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1589
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-467-920-18

Query Match 94.4%; Score 17; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 3.3e-12;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTFYCCELCCYPACAGC 17
Db 1 NTFYCCELCCYPACAGC 17

RESULT 12
US-08-467-920-23
; Sequence 23, Application US/08467920
; Patent No. 5962220
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; TITLE OF INVENTION: Compositions That Specifically
; TITLE OF INVENTION: Bind To Colorectal Cancer Cells
; TITLE OF INVENTION: And Methods Of Using The Same
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: No. 5962220ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,920
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/141,892
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1589
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FILING DATE:

```

US-08-467-920-23

Query Match 94.4%; Score 17; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3.3e-12;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TFYCCCLCCYPACAGCN 18  
|||||  
Db 1 TFYCCCLCCYPACAGCN 17

RESULT 13

US-08-635-930-18  
; Sequence 18, Application US/08635930  
; Patent No. 6060037

; GENERAL INFORMATION:

; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: Compositions That Specifically Bind To  
; TITLE OF INVENTION: Colorectal Cancer Cells And Methods Of Using  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6060037ris  
; STREET: One Liberty Place, 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19103

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: WINDOWS 3.1  
; SOFTWARE: WordPerfect 6.0/6.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/635,930  
; FILING DATE: 26-APR-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/141,892

; FILING DATE: 26-OCT-1993

; CLASSIFICATION: 435

; APPLICATION DATA:

; APPLICATION NUMBER: 08/305,056

; FILING DATE: 13-SEP-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Deluca, Mark

; REGISTRATION NUMBER: 33,229

; REFERENCE/DOCKET NUMBER: TJU-1360

; TELEPHONE: 215-568-3100

; TELEFAX: 215-568-3439

; INFORMATION FOR SEQ ID NO: 18:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 17 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-635-930-18

Query Match 94.4%; Score 17; DB 3; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3.3e-12;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTFYCCCLCCYPACAGC 17  
|||||  
Db 1 NTFYCCCLCCYPACAGC 17

RESULT 14

US-08-635-930-23  
; Sequence 23, Application US/08635930  
; Patent No. 6060037

; GENERAL INFORMATION:

; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: Compositions That Specifically Bind To  
; TITLE OF INVENTION: Colorectal Cancer Cells And Methods Of Using  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6060037ris  
; STREET: One Liberty Place, 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19103

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: WINDOWS 3.1  
; SOFTWARE: WordPerfect 6.0/6.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/635,930  
; FILING DATE: 26-APR-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/141,892

; FILING DATE: 26-OCT-1993

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/305,056

; FILING DATE: 13-SEP-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Deluca, Mark

; REGISTRATION NUMBER: 33,229

; REFERENCE/DOCKET NUMBER: TJU-1360

; TELEPHONE: 215-568-3100

; TELEFAX: 215-568-3439

; INFORMATION FOR SEQ ID NO: 23:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 17 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-635-930-23

Query Match 94.4%; Score 17; DB 3; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3.3e-12;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TFYCCCLCCYPACAGCN 18  
|||||  
Db 1 TFYCCCLCCYPACAGCN 17

RESULT 15

US-09-193-997-18  
; Sequence 18, Application US/09193997  
; Patent No. 6087109

; GENERAL INFORMATION:

; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: Compositions That Specifically  
; TITLE OF INVENTION: Bind To Colorectal Cancer Cells  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &amp;

; ADDRESS: No. 6087109ris

; STREET: One Liberty Place, 46th Floor

; CITY: Philadelphia

; STATE: Pennsylvania

; COUNTRY: USA

; ZIP: 19103

; COMPUTER READABLE FORM:



```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/193,997
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/467,920
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1589
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-193-997-18

```

```

Query Match          94.4%; Score 17; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 3.3e-12;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 NTFYCCCLCCYPACAGC 17
Db      1 NTFYCCCLCCYPACAGC 17

```

Search completed: March 26, 2005, 17:27:26  
Job time : 24.4286 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2005, 17:25:02 ; Search time 58.1786 Seconds  
(without alignments)  
102.440 Million cell updates/sec

Title: US-10-775-481A-3

Perfect score: 18

Sequence: 1 NTFYCCBLCYCPACAGCN 18

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 1407402 seqs, 331100923 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1407402

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pdb.p\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pdb.p\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pdb.p\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pdb.p\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pdb.p\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pdb.p\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pdb.p\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pdb.p\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pdb.p\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pdb.p\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pdb.p\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pdb.p\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pdb.p\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pdb.p\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pdb.p\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pdb.p\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pdb.p\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pdb.p\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pdb.p\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pdb.p\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 18    | 100.0       | 18     | 10    | US-09-930-915A-289 |
| 2          | 18    | 100.0       | 18     | 14    | US-10-082-014-79   |
| 3          | 18    | 100.0       | 18     | 14    | US-10-372-076-80   |
| 4          | 18    | 100.0       | 18     | 15    | US-10-621-684-3    |
| 5          | 18    | 100.0       | 18     | 16    | US-10-806-006-289  |
| 6          | 18    | 100.0       | 18     | 16    | US-10-677-074-80   |
| 7          | 18    | 100.0       | 18     | 16    | US-10-805-913-288  |
| 8          | 18    | 100.0       | 18     | 17    | US-10-796-719-3    |
| 9          | 17    | 94.4        | 17     | 15    | US-10-621-684-18   |
| 10         | 17    | 94.4        | 17     | 15    | US-10-621-684-23   |
| 11         | 16    | 88.9        | 16     | 15    | US-10-621-684-19   |
| 12         | 16    | 88.9        | 16     | 15    | US-10-621-684-24   |
| 13         | 15    | 83.3        | 15     | 15    | US-10-621-684-20   |

|    |    |      |    |    |                    |                    |
|----|----|------|----|----|--------------------|--------------------|
| 14 | 15 | 83.3 | 15 | 15 | US-10-621-684-25   | Sequence 25, Appl  |
| 15 | 15 | 83.3 | 18 | 10 | US-09-930-915A-290 | Sequence 290, Appl |
| 16 | 15 | 83.3 | 18 | 14 | US-10-082-014-80   | Sequence 80, Appl  |
| 17 | 15 | 83.3 | 18 | 14 | US-10-372-076-81   | Sequence 81, Appl  |
| 18 | 15 | 83.3 | 18 | 16 | US-10-806-006-290  | Sequence 290, Appl |
| 19 | 15 | 83.3 | 18 | 16 | US-10-677-074-81   | Sequence 81, Appl  |
| 20 | 15 | 83.3 | 18 | 16 | US-10-805-913-290  | Sequence 280, Appl |
| 21 | 14 | 77.8 | 14 | 14 | US-09-930-915A-288 | Sequence 288, Appl |
| 22 | 14 | 77.8 | 14 | 14 | US-10-082-014-78   | Sequence 78, Appl  |
| 23 | 14 | 77.8 | 14 | 14 | US-10-372-076-79   | Sequence 79, Appl  |
| 24 | 14 | 77.8 | 14 | 15 | US-10-621-684-21   | Sequence 21, Appl  |
| 25 | 14 | 77.8 | 14 | 15 | US-10-621-684-26   | Sequence 26, Appl  |
| 26 | 14 | 77.8 | 14 | 16 | US-10-806-006-288  | Sequence 288, Appl |
| 27 | 14 | 77.8 | 14 | 16 | US-10-677-074-79   | Sequence 79, Appl  |
| 28 | 14 | 77.8 | 14 | 16 | US-10-805-913-288  | Sequence 288, Appl |
| 29 | 13 | 72.2 | 13 | 15 | US-10-621-684-22   | Sequence 22, Appl  |
| 30 | 11 | 61.1 | 19 | 15 | US-10-621-684-49   | Sequence 49, Appl  |
| 31 | 10 | 55.6 | 17 | 15 | US-10-621-684-8    | Sequence 8, Appl   |
| 32 | 10 | 55.6 | 17 | 15 | US-10-621-684-41   | Sequence 41, Appl  |
| 33 | 10 | 55.6 | 18 | 15 | US-10-621-684-7    | Sequence 7, Appl   |
| 34 | 10 | 55.6 | 18 | 15 | US-10-621-684-13   | Sequence 13, Appl  |
| 35 | 10 | 55.6 | 18 | 15 | US-10-621-684-38   | Sequence 38, Appl  |
| 36 | 10 | 55.6 | 18 | 15 | US-10-621-684-39   | Sequence 39, Appl  |
| 37 | 10 | 55.6 | 18 | 15 | US-10-621-684-40   | Sequence 40, Appl  |
| 38 | 10 | 55.6 | 18 | 17 | US-10-796-719-2    | Sequence 2, Appl   |
| 39 | 10 | 55.6 | 18 | 17 | US-10-796-719-4    | Sequence 4, Appl   |
| 40 | 10 | 55.6 | 18 | 17 | US-10-796-719-11   | Sequence 11, Appl  |
| 41 | 10 | 55.6 | 19 | 15 | US-10-621-684-2    | Sequence 2, Appl   |
| 42 | 10 | 55.6 | 72 | 17 | US-10-796-719-20   | Sequence 20, Appl  |
| 43 | 9  | 50.0 | 16 | 15 | US-10-621-684-9    | Sequence 9, Appl   |
| 44 | 9  | 50.0 | 17 | 15 | US-10-621-684-14   | Sequence 14, Appl  |
| 45 | 8  | 44.4 | 15 | 15 | US-10-621-684-10   | Sequence 10, Appl  |

#### ALIGNMENTS

#### RESULT 1

US-09-930-915A-289  
; Sequence 289, Application US/09930915A  
; Publication No. US20030138769A1  
; GENERAL INFORMATION:  
; APPLICANT: Birkett, Ashley J.  
; TITLE OF INVENTION: IMMUNOGENIC HBC CHIMER PARTICLES HAVING ENHANCED STABILITY  
; FILE REFERENCE: 4564/83501 ICC-102.2 PCT  
; CURRENT APPLICATION NUMBER: US/09/930,915A  
; PRIOR FILING DATE: 2001-08-15  
; PRIOR APPLICATION NUMBER: 60/226,867  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,843  
; NUMBER OF SEQ ID NOS: 313  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 289  
; LENGTH: 18  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-09-930-915A-289

Query Match 100.0%; Score 18; DB 10; Length 18;  
Best Local Similarity 100.0%; Pred. No. 4.4e-13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTFYCCBLCYCPACAGCN 18

Db 1 NTFYCCBLCYCPACAGCN 18

#### RESULT 2

US-10-082-014-79  
; Sequence 79, Application US/10082014  
; Publication No. US20030185859A1

```
; GENERAL INFORMATION:
; APPLICANT: Birkett, Ashley J.
; TITLE OF INVENTION: IMMUNOGENIC HBC CHIMER PARTICLES STABILIZED WITH AN N-TERMINAL CY
; FILE REFERENCE: ICC-130.0 4564/85124
; CURRENT APPLICATION NUMBER: US/10/082,014
; CURRENT FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 09/930,915
; PRIOR FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 290
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 79
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-082-014-79

Query Match 100.0%; Score 18; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.4e-13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTFYCCELCCYPACAGCN 18
Db 1 NTFYCCELCCYPACAGCN 18
|||||

RESULT 3
US-10-372-076-80
; Sequence 80, Application US/10372076
; Publication No. US20030198645A1
; GENERAL INFORMATION:
; APPLICANT: Page, Mark
; APPLICANT: Friede, Martin
; TITLE OF INVENTION: STABILIZED HBC CHIMER PARTICLES AS THERAPEUTIC VACCINE FOR
; TITLE OF INVENTION: CHRONIC HEPATITIS
; FILE REFERENCE: 4564/87179
; CURRENT APPLICATION NUMBER: US/10/372,076
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: 10/080,299
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 10/082,014
; PRIOR FILING DATE: 2002-02-22
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 80
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-372-076-80

Query Match 100.0%; Score 18; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.4e-13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTFYCCELCCYPACAGCN 18
Db 1 NTFYCCELCCYPACAGCN 18
|||||

RESULT 4
US-10-621-684-3
; Sequence 3, Application US/10621684
; Publication No. US20040029182A1
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; TITLE OF INVENTION: St Receptor Binding Compounds and
; TITLE OF INVENTION: Methods of Using the Same
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. US20040029182A1Iris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA

Query Match 100.0%; Score 18; DB 16; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.4e-13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTFYCCELCCYPACAGCN 18
Db 1 NTFYCCELCCYPACAGCN 18
|||||
```

```
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/621,684
; FILING DATE: 17-Jul-2003
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,447A
; FILING DATE: 05-JAN-1996
; APPLICATION NUMBER: US 08/141,892
; FILING DATE: 26-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1702
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-621-684-3

Query Match 100.0%; Score 18; DB 15; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.4e-13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTFYCCELCCYPACAGCN 18
Db 1 NTFYCCELCCYPACAGCN 18
|||||

RESULT 5
US-10-806-006-289
; Sequence 289, Application US/10806006
; Publication No. US20040152876A1
; GENERAL INFORMATION:
; APPLICANT: Birkett, Ashley J.
; TITLE OF INVENTION: IMMUNOGENIC HBC CHIMER PARTICLES HAVING ENHANCED
; TITLE OF INVENTION: STABILITY
; FILE REFERENCE: 4564/91644 ICC-102.2 DV I
; CURRENT APPLICATION NUMBER: US/10/806,006
; CURRENT FILING DATE: 2004-03-22
; PRIOR APPLICATION NUMBER: 09/930,915
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: PCT/US01/41759
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60/226,867
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,843
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 313
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 289
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-806-006-289

Query Match 100.0%; Score 18; DB 16; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.4e-13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTFYCCELCCYPACAGCN 18
Db 1 NTFYCCELCCYPACAGCN 18
|||||
```

Db 1 NTFYCCCLCCYPACAGCN 18

## RESULT 6

US-10-677-074-80  
; Sequence 80, Application US/10677074  
; Publication No. US20040156863A1  
; GENERAL INFORMATION:  
; APPLICANT: Page, Mark  
; APPLICANT: Friede, Martin  
; APPLICANT: Schmidt, Annette Elisabeth  
; APPLICANT: Stober, Detlef  
; TITLE OF INVENTION: STABILIZED HBC CHIMER PARTICLES AS THERAPEUTIC VACCINE FOR  
; TITLE OF INVENTION: CHRONIC HEPATITIS  
; FILE REFERENCE: 4564/87179  
; CURRENT APPLICATION NUMBER: US/10/677,074  
; CURRENT FILING DATE: 2003-10-01  
; PRIOR APPLICATION NUMBER: 10/372,076  
; PRIOR FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: 10/080,299  
; PRIOR FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: 10/082,014  
; PRIOR FILING DATE: 2002-02-22  
; NUMBER OF SEQ ID NOS: 308  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 80  
; LENGTH: 18  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-10-677-074-80

Query Match 100.0%; Score 18; DB 16; Length 18;  
Best Local Similarity 100.0%; Pred. No. 4.4e-13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTFYCCCLCCYPACAGCN 18

Db 1 NTFYCCCLCCYPACAGCN 18

## RESULT 7

US-10-805-913-289  
; Sequence 289, Application US/10805913  
; Publication No. US20040156864A1  
; GENERAL INFORMATION:  
; APPLICANT: Birkett, Ashley J.  
; TITLE OF INVENTION: IMMUNOGENIC HBC CHIMER PARTICLES HAVING ENHANCED  
; TITLE OF INVENTION: STABILITY  
; FILE REFERENCE: 4564/91645 ICC-102.2 DV II  
; CURRENT APPLICATION NUMBER: US/10/805,913  
; CURRENT FILING DATE: 2004-03-22  
; PRIOR APPLICATION NUMBER: 09/930,915  
; PRIOR FILING DATE: 2001-08-15  
; PRIOR APPLICATION NUMBER: PCT/US01/41759  
; PRIOR FILING DATE: 2001-08-16  
; PRIOR APPLICATION NUMBER: 60/226,867  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,843  
; PRIOR FILING DATE: 2000-08-16  
; NUMBER OF SEQ ID NOS: 313  
; SOFTWARE: PatentIn ver. 2.1  
; SEQ ID NO 289  
; LENGTH: 18  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-10-805-913-289

Query Match 100.0%; Score 18; DB 16; Length 18;  
Best Local Similarity 100.0%; Pred. No. 4.4e-13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTFYCCCLCCYPACAGCN 18

|||||

Db 1 NTFYCCCLCCYPACAGCN 18

## RESULT 8

US-10-796-719-3  
; Sequence 3, Application US/10796719  
; Publication No. US20050020811A1  
; GENERAL INFORMATION:  
; APPLICANT: Currie, Mark G.  
; APPLICANT: Mahajan-Miklos, Shalina  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE  
; TITLE OF INVENTION: TREATMENT OF GASTROINTESTINAL DISORDERS  
; FILE REFERENCE: 14184-043001  
; CURRENT APPLICATION NUMBER: US/10/796,719  
; CURRENT FILING DATE: 2004-03-09  
; PRIOR APPLICATION NUMBER: US 10/766,735  
; PRIOR FILING DATE: 2004-01-28  
; PRIOR APPLICATION NUMBER: US 60/443,098  
; PRIOR FILING DATE: 2003-01-28  
; PRIOR APPLICATION NUMBER: US 60/471,288  
; PRIOR FILING DATE: 2003-05-15  
; PRIOR APPLICATION NUMBER: US 60/519,460  
; PRIOR FILING DATE: 2003-11-12  
; NUMBER OF SEQ ID NOS: 149  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 18  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-10-796-719-3

Query Match 100.0%; Score 18; DB 17; Length 18;  
Best Local Similarity 100.0%; Pred. No. 4.4e-13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTFYCCCLCCYPACAGCN 18

Db 1 NTFYCCCLCCYPACAGCN 18

## RESULT 9

US-10-621-684-18  
; Sequence 18, Application US/10621684  
; Publication No. US20040029182A1  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: ST Receptor Binding Compounds and  
; METHODS OF USING THE SAME  
; NUMBER OF SEQUENCES: 56  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. US20040029182A1 is  
; STREET: One Liberty Place, 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: WordPerfect 6.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/621,684  
; FILING DATE: 17-Jul-2003  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/583,447A  
; FILING DATE: 05-JAN-1996  
; APPLICATION NUMBER: US 08/141,892  
; FILING DATE: 26-OCT-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229

```

; REFERENCE/DOCKET NUMBER: TJU-1702
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-10-621-684-18
      94.4%; Score 17; DB 15; Length 17;
Query Match 100.0%; Pred. No. 5.4e-12; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 0;

Qy 1 NTFYCCCLCCYPACAGC 17
      |||||
Db 1 NTFYCCCLCCYPACAGC 17

RESULT 10
US-10-621-684-23
; Sequence 23, Application US/10621684
; Publication No. US20040029182A1
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; TITLE OF INVENTION: ST Receptor Binding Compounds and
; Methods of Using the Same
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. US20040029182A1ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/621,684
; FILING DATE: 17-Jul-2003
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,447A
; FILING DATE: 05-JAN-1996
; APPLICATION NUMBER: US 08/141,892
; FILING DATE: 26-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1702
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-10-621-684-19
      Query Match 88.9%; Score 16; DB 15; Length 16;
      Best Local Similarity 100.0%; Pred. No. 6.5e-11;
      Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TFYCCCLCCYPACAGC 17
      |||||
Db 1 TFYCCCLCCYPACAGC 16

RESULT 12
US-10-621-684-24
; Sequence 24, Application US/10621684
; Publication No. US20040029182A1
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; TITLE OF INVENTION: ST Receptor Binding Compounds and
; Methods of Using the Same
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. US20040029182A1ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/621,684
; FILING DATE: 17-Jul-2003
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,447A
; FILING DATE: 05-JAN-1996
; APPLICATION NUMBER: US 08/141,892
; FILING DATE: 26-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1702
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-10-621-684-23
      Query Match 94.4%; Score 17; DB 15; Length 17;
      Best Local Similarity 100.0%; Pred. No. 5.4e-12; Indels 0; Gaps 0;
      Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TFYCCCLCCYPACAGC 18
      |||||
Db 1 TFYCCCLCCYPACAGC 18
```

```

;
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1702
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-10-621-684-20

Query Match 83.3%; Score 15; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 7.9e-10;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 FYCCELCCYPACAGC 17
DB 1 FYCCELCCYPACAGC 15

RESULT 14
US-10-621-684-25
; Sequence 25, Application US/10621684
; Publication No. US20040029182A1
; GENERAL INFORMATION:
; APPLICANT: Waldman, Scott A.
; TITLE OF INVENTION: ST Receptor Binding Compounds and
; METHODS OF USING THE SAME
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Woodcock Washburn Kurtz Mackiewicz & No. US20040029182A1ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/621,684
; FILING DATE: 17-Jul-2003
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,447A
; FILING DATE: 05-JAN-1996
; APPLICATION NUMBER: US 08/141,892
; FILING DATE: 26-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1702
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-10-621-684-25

Query Match 83.3%; Score 15; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 7.9e-10;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 4 YCCELCGYPACGCN 18  
 |||||  
 Db 1 YCCELCGYPACGCN 15

RESULT 15

US-09-930-915A-290  
 ; Sequence 290, Application US/09930915A  
 ; Publication No. US20030138769A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Birkett, Ashley J.  
 ; TITLE OF INVENTION: IMMUNOGENIC HBC CHIMER PARTICLES HAVING ENHANCED  
 ; FILE OF INVENTION: STABILITY  
 ; FILE REFERENCE: 4564/83501 ICC-102.2 PCT  
 ; CURRENT APPLICATION NUMBER: US/09/930,915A  
 ; CURRENT FILING DATE: 2001-08-15  
 ; PRIOR APPLICATION NUMBER: 60/226,867  
 ; PRIOR FILING DATE: 2000-08-22  
 ; PRIOR APPLICATION NUMBER: 60/225,843  
 ; PRIOR FILING DATE: 2000-08-16  
 ; NUMBER OF SEQ ID NOS: 313  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 290  
 ; LENGTH: 18  
 ; TYPE: PRT  
 ; ORGANISM: Escherichia coli  
 US-09-930-915A-290

Query Match 83.3%; Score 15; DB 10; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 9e-10;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 YCCELCGYPACGCN 18  
 |||||  
 Db 4 YCCELCGYPACGCN 18

Search completed: March 26, 2005, 17:44:31  
 Job time : 59.1786 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2005, 17:10:31 ; Search time 18.6429 Seconds  
(without alignments)  
92.899 Million cell updates/sec

Title: US-10-775-481A-3  
Perfect score: 18  
Sequence: 1 NTFYCCBLCCYPACAGCN 18  
Scoring table: OMIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues  
Word size : 0  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Listing first 45 summaries

Database : PIR 79.\*  
1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 18    | 100.0       | 18     | 1 QHEC2  | heat-stable entero |
| 2          | 10    | 55.6        | 18     | 2 A60103 | heat-stable entero |
| 3          | 10    | 55.6        | 72     | 1 QHEC1  | heat-stable entero |
| 4          | 7     | 38.9        | 72     | 1 QHEC4  | heat-stable entero |
| 5          | 7     | 38.9        | 72     | 1 QHEC1B | heat-stable entero |
| 6          | 6     | 33.3        | 53     | 2 S68705 | heat-stable entero |
| 7          | 6     | 33.3        | 66     | 2 S31652 | enterotoxin - Vers |
| 8          | 6     | 33.3        | 71     | 2 S25659 | heat-stable entero |
| 9          | 6     | 33.3        | 90     | 2 D85845 | probable lysis pro |
| 10         | 6     | 33.3        | 534    | 2 T15414 | hypothetical prote |
| 11         | 6     | 33.3        | 862    | 2 T29133 | hypothetical prote |
| 12         | 6     | 33.3        | 924    | 2 E71476 | alanine-tRNA ligas |
| 13         | 6     | 33.3        | 2482   | 2 I48922 | cation-independent |
| 14         | 6     | 33.3        | 2483   | 1 A49617 | insulin-like growt |
| 15         | 5     | 27.8        | 72     | 2 C39370 | rhombotin homolog  |
| 16         | 5     | 27.8        | 74     | 2 AF0925 | hypothetical prote |
| 17         | 5     | 27.8        | 75     | 2 H84631 | hypothetical prote |
| 18         | 5     | 27.8        | 80     | 2 A82531 | hypothetical prote |
| 19         | 5     | 27.8        | 108    | 2 AF1043 | hypothetical prote |
| 20         | 5     | 27.8        | 111    | 2 I74319 | gene BHS-2 protein |
| 21         | 5     | 27.8        | 112    | 2 B72476 | hypothetical prote |
| 22         | 5     | 27.8        | 129    | 2 A24255 | chorion class A pr |
| 23         | 5     | 27.8        | 133    | 2 T17300 | hypothetical prote |
| 24         | 5     | 27.8        | 135    | 2 S34815 | nifU protein - Rh  |
| 25         | 5     | 27.8        | 156    | 2 A32795 | T-cell translocati |
| 26         | 5     | 27.8        | 187    | 2 T17464 | probable rRNA meth |
| 27         | 5     | 27.8        | 194    | 2 A83103 | conserved hypotet  |
| 28         | 5     | 27.8        | 224    | 2 F81726 | conserved hypotet  |
| 29         | 5     | 27.8        | 228    | 2 D95865 | probable pentose-5 |

|    |   |      |     |   |        |                     |
|----|---|------|-----|---|--------|---------------------|
| 30 | 5 | 27.8 | 230 | 2 | T20398 | hypothetical prote  |
| 31 | 5 | 27.8 | 235 | 2 | F70405 | hypothetical prote  |
| 32 | 5 | 27.8 | 243 | 2 | D85888 | hypothetical prote  |
| 33 | 5 | 27.8 | 243 | 2 | A91044 | hypothetical prote  |
| 34 | 5 | 27.8 | 246 | 2 | AG2988 | competence protein  |
| 35 | 5 | 27.8 | 247 | 2 | AC0923 | probable uridine p  |
| 36 | 5 | 27.8 | 263 | 2 | G87721 | protein ZC123.3 [   |
| 37 | 5 | 27.8 | 265 | 2 | AG0814 | cob(II)alamin adeno |
| 38 | 5 | 27.8 | 267 | 2 | B65021 | cob(II)alamin adeno |
| 39 | 5 | 27.8 | 275 | 2 | AC0189 | probable exported   |
| 40 | 5 | 27.8 | 277 | 1 | QOBE2  | HKRF protein - hu   |
| 41 | 5 | 27.8 | 291 | 2 | A9295  | competence protein  |
| 42 | 5 | 27.8 | 306 | 2 | T24402 | hypothetical prote  |
| 43 | 5 | 27.8 | 308 | 2 | E90179 | hypothetical prote  |
| 44 | 5 | 27.8 | 316 | 1 | D71246 | hypothetical prote  |
| 45 | 5 | 27.8 | 316 | 2 | T49224 | SRG1-like protein   |

ALIGNMENTS

RESULT 1

QHEC2  
heat-stable enterotoxin ST-2 - Escherichia coli  
C:Species: Escherichia coli  
C:Date: 06-Jul-1982 #sequence\_revision 06-Jul-1982 #text\_change 09-Jul-2004  
C:Accession: A01823  
R:Chan, S.K.; Giannella, R.A.  
J. Biol. Chem. 256, 7744-7746, 1981  
A:Title: Amino acid sequence of heat-stable enterotoxin produced by Escherichia coli pat  
A:Reference number: A01823; MUID:81264141; PMID:7021541  
A:Accession: A01823  
A:Molecule type: protein  
A:Residues: 1-18 <CHA>  
A:Cross-references: UNIPROT:P01560  
A:Experimental source: strain 18D, serotype 0.42:k86:H37  
C:Comment: This enterotoxin is one of several, of differing molecular sizes, produced by  
idues of the heat-stable enterotoxin ST-1.  
C:Superfamily: heat-stable enterotoxin ST  
C:Keywords: enterotoxin; heat-stable protein  
F:1-18/Product: heat-stable enterotoxin ST-2 #status experimental <MAT>  
F:5-10,6-14,9-17/Disulfide bonds: #status predicted

Query Match 100.0%; Score 18; DB 1; Length 18;  
Best Local Similarity 100.0%; Pred. No. 2.7e-15;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTFYCCBLCCYPACAGCN 18  
|||  
Db 1 NTFYCCBLCCYPACAGCN 18

RESULT 2

A60103  
heat-stable enterotoxin ST-Ia - Citrobacter freundii  
C:Species: Citrobacter freundii  
C:Date: 10-Nov-1992 #sequence\_revision 10-Nov-1992 #text\_change 09-Jul-2004  
C:Accession: A60103  
R:Guarino, A.; Giannella, R.; Thompson, M.R.  
Infect. Immun. 57, 649-652, 1989  
A:Title: Citrobacter freundii produces an 18-amino-acid heat-stable enterotoxin identical  
A:Reference number: A60103; MUID:89108617; PMID:2912902  
A:Accession: A60103  
A:Molecule type: protein  
A:Residues: 1-18 <GUA>  
A:Cross-references: UNIPROT:Q7M0U3  
C:Superfamily: heat-stable enterotoxin ST

Query Match 55.6%; Score 10; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 1.2e-05;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTFYCCBLCC 10

Db 1 NTFYCCCLCC 10  
|||||  
RESULT 3  
QHEC1  
heat-stable enterotoxin ST-I precursor - Escherichia coli  
N:Alternate names: heat-stable enterotoxin estA1  
C:Species: Escherichia coli  
C:Date: 31-Aug-1980 #sequence\_revision 31-Aug-1980 #text\_change 09-Jul-2004  
C:Accession: A01822; A30985; A36732; JT0374; I51932  
R:So, M.; McCarthy, B.J.  
Proc. Natl. Acad. Sci. U.S.A. 77, 4011-4015, 1980  
A:Title: Nucleotide sequence of the bacterial transposon Tni681 encoding a heat-stable (S)  
A:Reference number: A01822; MUID:81054703; PMID:6254008  
A:Accession: A01822  
A:Molecule type: DNA  
A:Residues: 1-72 <LAZ>  
A:Cross-references: UNIPROT:P01559; GB:V00612; GB:J01831; NID:G43704; PIDN:CAA23883.1; R  
R:Lazure, C.; Seidah, N.G.; Chretien, M.; Lallier, R.; St-Pierre, S.  
Can. J. Biochem. Cell Biol. 61, 287-292, 1983  
A:Title: Primary structure determination of Escherichia coli heat-stable enterotoxin of  
A:Reference number: A30985; MUID:83284515; PMID:6349752  
A:Accession: A30985  
A:Molecule type: protein  
A:Residues: 55-72 <LAZ>  
A:Experimental source: strain F11  
R:Dallas, W.S.  
J. Bacteriol. 172, 5490-5493, 1990  
A:Title: The heat-stable toxin I gene from Escherichia coli 18D.  
A:Reference number: A36732; MUID:90368614; PMID:2203756  
A:Accession: A36732  
A:Molecule type: DNA  
A:Residues: 1-72 <DAL>  
A:Cross-references: GB:M58746; NID:G145860; PIDN:AAA62776.1; PID:G145861  
A:Experimental source: strain 18D  
R:Stieglitz, H.; Cervantes, L.; Robledo, R.; Fonseca, R.; Covarrubias, L.; Bolivar, F.;  
Plasmid 20, 42-53, 1988  
A:Title: Cloning, sequencing, and expression in ficoll-generated minicells of an Escheri  
A:Reference number: JT0373; MUID:89202548; PMID:3071819  
A:Accession: JT0374  
A:Molecule type: DNA  
A:Residues: 1-72 <STI>  
R:Seikizaki, T.; Akashi, H.; Terakado, N.  
Am. J. Vet. Res. 46, 909-912, 1985  
A:Title: Nucleotide sequences of the genes for Escherichia coli heat-stable enterotoxin  
A:Reference number: I51932; MUID:85249571; PMID:2990268  
A:Accession: I51932  
A:Status: translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-69, 'P', 71-72 <RES>  
A:Cross-references: GB:M25607; NID:G147877; PIDN:AAA24653.1; PID:G147878  
C:Comment: Both heat-stable and heat-labile enterotoxins are produced by pathogenic stra  
ular sizes.  
C:Superfamily: heat-stable enterotoxin ST  
C:Keywords: enterotoxin; heat-stable protein  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-54/Domain: propeptide #status predicted <PRO>  
F:55-72/Product: heat-stable enterotoxin ST-I #status experimental <MAT>  
F:59-64,60-68,63-71/Disulfide bonds: #status predicted  
Query Match 55.6%; Score 10; DB 1; Length 72;  
Best Local Similarity 100.0%; Pred. No. 3e-05; Mismatches 0; Indels 0; Gaps 0;  
Matches 10; Conservative 0

QY 1 NTFYCCCLCC 10  
|||||  
Db 55 NTFYCCCLCC 64  
|||||

RESULT 4  
QHEC4  
heat-stable enterotoxin STA4 precursor - Escherichia coli

C:Species: Escherichia coli  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 09-Jul-2004  
C:Accession: JT0373; A35978  
R:Stieglitz, H.; Cervantes, L.; Robledo, R.; Fonseca, R.; Covarrubias, L.; Bolivar, F.;  
Plasmid 20, 42-53, 1988  
A:Title: Cloning, sequencing, and expression in ficoll-generated minicells of an Escheric  
A:Reference number: JT0373; MUID:89202548; PMID:3071819  
A:Accession: JT0373  
A:Molecule type: DNA  
A:Residues: 1-72 <STI>  
A:Cross-references: UNIPROT:P07965; GB:J03311; NID:G147875; PIDN:AAA24652.1; PID:G147876  
Toxicon 28, 453-456, 1990  
R:Zhou, X.; Shen, L.P.; Chi, C.W.  
A:Title: Isolation and nucleotide sequence determination of a gene encoding a heat-stabl  
A:Reference number: A35978; MUID:90273381; PMID:2190361  
A:Accession: A35978  
A:Molecule type: DNA  
A:Residues: 1-72 <ZHO>  
C:Genetics:  
A:Gene: estA4  
C:Superfamily: heat-stable enterotoxin ST  
C:Keywords: enterotoxin; heat-stable protein  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-53/Domain: propeptide #status predicted <PRO>  
F:54-72/Product: heat-stable enterotoxin #status predicted <MAT>  
F:59-64,60-68,63-71/Disulfide bonds: #status predicted  
Query Match 38.9%; Score 7; DB 1; Length 72;  
Best Local Similarity 100.0%; Pred. No. 0.13; Mismatches 0; Indels 0; Gaps 0;  
Matches 7; Conservative 0

QY 4 YCCCLCC 10  
|||||  
Db 58 YCCCLCC 64  
|||||

RESULT 5  
QHECIB  
heat-stable enterotoxin ST-Ib precursor - Escherichia coli  
N:Alternate names: heat-stable enterotoxin ST-A2  
C:Species: Escherichia coli  
C:Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 09-Jul-2004  
C:Accession: JS0292; A33068; A33067; A30567  
R:Moseley, S.L.; Hardy, J.W.; Hug, M.I.; Echeverria, P.; Falkow, S.  
Infect. Immun. 39, 1167-1174, 1983  
A:Title: Isolation and nucleotide determination of a gene encoding a heat-stable enteroto  
A:Reference number: JS0292; MUID:83184648; PMID:6341230  
A:Accession: JS0292  
A:Molecule type: DNA  
A:Residues: 1-72 <MOS>  
A:Cross-references: UNIPROT:Q47185; UNIPROT:P07965; GB:M34916; NID:G146407; PIDN:AAA2399(C  
R:Dwarakanath, P.; Visweswariah, S.S.; Subrahmanyam, Y.V.B.K.; Shanthi, G.; Jagannatha, P  
Gene 81, 219-226, 1989  
A:Title: Cloning and hyperexpression of a gene encoding the heat-stable toxin of Escheric  
A:Reference number: A33068; MUID:90034194; PMID:2680769  
A:Accession: A33068  
A:Molecule type: DNA  
A:Residues: 1-18, 'A', 20-72 <DWA>  
A:Cross-references: GB:M23255; NID:G148029; PIDN:AAA24686.1; PID:G148030  
R:Aimoto, S.; Takao, T.; Shimonishi, Y.; Hara, S.; Takeda, T.; Takeda, Y.; Miwatani, T.  
Eur. J. Biochem. 129, 257-263, 1982  
A:Title: Amino acid sequence of heat-stable enterotoxin produced by human enterotoxigenic  
A:Reference number: A33067; MUID:83105138; PMID:6759126  
A:Accession: A33067  
A:Molecule type: protein  
A:Residues: 54-72 <AIM>  
R:Guzman-Verduzco, L.M.; Kupersztovich, Y.M.  
Infect. Immun. 57, 645-648, 1989  
A:Title: Rectification of two Escherichia coli heat-stable enterotoxin allele sequences  
A:Reference number: A30567; MUID:89108616; PMID:2643580  
A:Accession: A30567  
A:Molecule type: DNA

A;Residues: 1-18,'A',20-24,'AG',27-41,'V',43-44,'N',46,'E',48,'S',50-72 <GUZ>  
A;Cross-references: GB:M18345; NID:g145862; PIDN:AAA23729.1; PID:g145863  
C;Comment: This is one of the type I heat-stable enterotoxins that are methanol-soluble.  
C;Genetics:  
A;Gene: st  
C;Superfamily: heat-stable enterotoxin ST  
C;Keywords: enterotoxin; heat-stable protein  
F;1-53/Domain: signal sequence and propeptide #status predicted <SIG>  
F;54-72/Product: heat-stable enterotoxin ST-1b #status experimental <MAT>  
F;59-64,60-68,63-71/Disulfide bonds: #status experimental

Query Match 38.9%; Score 7; DB 1; Length 72;  
Best Local Similarity 100.0%; Pred. No. 0.13; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0

Qy 4 YCCCLCC 10  
Db 58 YCCCLCC 64

RESULT 6  
S31652  
heat-stable enterotoxin Y-STC - Yersinia enterocolitica  
C;Species: Yersinia enterocolitica  
C;Date: 25-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 13-Mar-1998  
C;Accession: S68705  
R;Yoshino, K.; Takao, T.; Huang, X.; Murata, H.; Nakao, H.; Takeda, T.; Shimonishi, Y.  
FEBS Lett. 362, 319-322, 1995  
A;Title: Characterization of a highly toxic, large molecular size heat-stable enterotoxin  
A;Reference number: S68705; MUID:95246844; PMID:7729521  
A;Accession: S68705  
A;Molecule type: protein  
A;Residues: 1-53 <YOS>  
A;Experimental source: strain 86-11  
C;Superfamily: heat-stable enterotoxin ST  
C;Keywords: enterotoxin; heat-stable protein  
F;41-46,42-50,45-53/Disulfide bonds: #status predicted

Query Match 33.3%; Score 6; DB 2; Length 53;  
Best Local Similarity 100.0%; Pred. No. 1.7; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0

Qy 5 CCECLCC 10  
Db 41 CCECLCC 46

RESULT 7  
S31652  
enterotoxin - Yersinia kristensenii  
C;Species: Yersinia kristensenii  
C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004  
C;Accession: S31652  
R;Ibrahim, A.; Liesack, W.; Stackebrandt, E.  
submitted to the EMBL Data Library, November 1992  
A;Reference number: S31652  
A;Accession: S31652  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-56 <IBR>  
A;Cross-references: UNIPROT:P31518; EMBL:X69218; NID:g48617; PIDN:CAA49152.1; PID:g48618  
C;Superfamily: heat-stable enterotoxin ST

Query Match 33.3%; Score 6; DB 2; Length 66;  
Best Local Similarity 100.0%; Pred. No. 1.9; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0

Qy 12 PACAGC 17  
Db 61 PACAGC 66

RESULT 8

S25659  
heat-stable enterotoxin yst precursor - Yersinia enterocolitica  
C;Species: Yersinia enterocolitica  
C;Date: 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 09-Jul-2004  
C;Accession: S25659; A41474; A23114; S65849  
R;Ibrahim, A.; Liesack, W.; Pike, S.; Stackebrandt, E.  
FEMS Microbiol. Lett. 97, 63-66, 1992  
A;Title: The polymerase chain reaction: an epidemiological tool to differentiate between  
F;54-72/Product: heat-stable enterotoxin ST-1b #status experimental <MAT>  
A;Reference number: S25659  
A;Accession: S25659  
A;Status: preliminary; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-71 <IBR>  
A;Cross-references: UNIPROT:P07593; EMBL:X65999; NID:g48611; PIDN:CAA46801.1; PID:g48612  
R;Delor, I.; Kaeckenbeeck, A.; Wauters, G.; Cornelis, G.R.  
Infect. Immun. 58, 2983-2988, 1990  
A;Title: Nucleotide sequence of yst, the Yersinia enterocolitica gene encoding the heat-  
A;Reference number: A41474; MUID:90354067; PMID:2201642  
A;Accession: A41474  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 1-47,'S',49-71 <DEL>  
A;Cross-references: GB:U09235; NID:g487394; PIDN:AAA18472.1; PID:g487395  
R;Takao, T.; Tominaga, N.; Yoshimura, S.; Shimonishi, Y.; Hara, S.; Inoue, T.; Miyama, A.  
Eur. J. Biochem. 152, 199-206, 1985  
A;Title: Isolation, primary structure and synthesis of heat-stable enterotoxin produced  
A;Reference number: A23114; MUID:86004705; PMID:4043080  
A;Accession: A23114  
A;Molecule type: protein  
A;Residues: 54-71 <YAK>  
R;Mikulskis, A.V.; Delor, I.; Ha Thi, V.; Cornelis, G.R.  
Mol. Microbiol. 14, 905-915, 1994  
A;Title: Regulation of the Yersinia enterocolitica enterotoxin yst gene. Influence of gr  
A;Reference number: S65849; MUID:95231297; PMID:7715452  
A;Accession: S65849  
A;Status: preliminary; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-43 <MIK>  
A;Cross-references: EMBL:U09235  
C;Genetics:  
A;Gene: yst  
C;Superfamily: heat-stable enterotoxin ST  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-41/Domain: propeptide #status predicted <PRO>  
F;42-71/Product: heat-stable enterotoxin yst #status predicted <MAT>

Query Match 33.3%; Score 6; DB 2; Length 71;  
Best Local Similarity 100.0%; Pred. No. 2; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0

Qy 12 PACAGC 17  
Db 66 PACAGC 71

RESULT 9  
D85845  
probable lysis protein S of prophage CP-933V Z3340 [imported] - Escherichia coli (strain  
C;Species: Escherichia coli  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C;Accession: D85845  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimailanta, E.; Potamouis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A;Reference number: A85480; MUID:21074935; PMID:11206551  
A;Accession: D85845  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-90 <STO>  
A;Cross-references: UNIPROT:Q8X4M8; GB:AE005174; NID:g12516391; PIDN:AGS7224.1; GSPDB:G  
A;Experimental source: strain O157:H7, substrain EDL933  
C;Genetics:

A:Gene: Z3340

Query Match 33.3%; Score 6; DB 2; Length 90;  
Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 FYCCEL 8

DB 7 FYCCEL 12

RESULT 10

T15414

hypothetical protein C05E11.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004

C:Accession: T15414

R:Geisel, C.

A:Description: The sequence of C. elegans cosmid C05E11.

A:Reference number: Z18347

A:Accession: T15414

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-534 <GEI>

A:Cross-references: UNIPROT:P54145; EMBL:U53338; NID:g1255840; PID:g1255843; PIDN:AAA961

A:Experimental source: strain Bristol N2; clone C05E11

C:Genetics:

A:Gene: CESP:C05E11.4

A:Map position: X

A:Introns: 43/1; 151/1; 182/3; 229/3; 339/2; 394/3; 423/3; 490/1

C:Superfamily: ammonium transport protein amt1

Query Match

Best Local Similarity 33.3%; Score 6; DB 2; Length 534;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 ACAGCN 18

DB 307 ACAGCN 312

RESULT 11

T29133

hypothetical protein SCIF2.24 - Streptomyces coelicolor (fragment)

C:Species: Streptomyces coelicolor

C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T29133

R:Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

A:Submitted to the EMBL Data Library, August 1998

A:Reference number: Z17215

A:Accession: T29133

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-862 <PAR>

A:Cross-references: UNIPROT:Q8CJM2; EMBL:AL031350; NID:e1316892; PID:e1316916; PIDN:CAA2

C:Genetics:

A:Note: SCIF2.24

Query Match

Best Local Similarity 33.3%; Score 6; DB 2; Length 862;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 YPACAG 16

DB 397 YPACAG 402

RESULT 12

E71476

alanine-tRNA ligase (EC 6.1.1.7) - Chlamydia trachomatis (serotype D, strain UW3/Cx)

C:Species: Chlamydia trachomatis

C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 17-Mar-2003

C:Accession: E71476  
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998  
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tract  
A:Reference number: A71570; MUID:99000809; PMID:9784136  
A:Accession: E71476  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-924 <ARN>  
A:Cross-references: GB:AE001346; GB:AE001273; NID:g3329203; PIDN:AAC68344.1; PID:g3329203  
A:Experimental source: serotype D, strain UW-3/Cx  
C:Genetics:  
A:Gene: alas  
C:Superfamily: alanyl-tRNA ligase  
C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 33.3%; Score 6; DB 2; Length 924;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 FYCCEL 8

DB 39 FYCCEL 44

RESULT 13

I48922

cation-independent mannose 6-phosphate/insulin-like growth factor II receptor precursor

C:Species: Mus musculus (house mouse)

C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 20-Aug-1999

C:Accession: I48922

R:Ludwig, T.; Tenschler, K.; Remmler, J.; Hoflack, B.; Lobel, P.

Gene 142, 311-312, 1994

A:Title: Cloning and sequencing of cDNAs encoding the full length mouse mannose 6-phosphate

A:Reference number: I48922; MUID:94252588; PMID:8194771

A:Accession: I48922

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-2482 <RES>

A:Cross-references: EMBL:U04710; NID:g451552; PIDN:AAA19568.1; PID:g451553

C:Superfamily: mannose 6-phosphate receptor, cation-independent; fibronectin type II repeat

C:Keywords: growth factor receptor

F:1895-1934/Domain: fibronectin type II repeat homology <2F1>

Query Match

Best Local Similarity 33.3%; Score 6; DB 2; Length 2482;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 YPACAG 16

DB 2177 YPACAG 2182

RESULT 14

A49617

insulin-like growth factor II / cation-independent mannose 6-phosphate receptor precursor

C:Species: Mus musculus (house mouse)

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004

C:Accession: A49617; B39538; A48700; A46439; I52815; I65739; A57529

R:Szebenyi, G.; Rotwein, P.

Genomics 19, 120-129, 1994

A:Title: The mouse insulin-like growth factor II/cation-independent mannose 6-phosphate

A:Reference number: A49617; MUID:94245146; PMID:8188212

A:Accession: A49617

A:Molecule type: DNA

A:Residues: 1-2483 <SZE>

A:Cross-references: UNIPROT:Q07113; GB:I22143; NID:g431410; PIDN:AAA39320.1; PID:g431412

R:Szebenyi, G.; Rotwein, P.

J. Biol. Chem. 266, 5534-5539, 1991

A:Title: Differential regulation of mannose 6-phosphate receptors and their ligands durin

A:Reference number: A39538; MUID:91170218; PMID:1848553

A:Accession: B39538

A:Molecule type: DNA

Proc. Natl. Acad. Sci. U.S.A. 88, 4367-4371, 1991  
A;Title: The thrombotin family of cysteine-rich LIM-domain oncogenes: distinct members ar  
A;Reference number: A39370; MUID:91239569; PMID:2034676  
A;Accession: C39370  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-72 <BOE>  
A;Cross-references: GB:M64358  
C;Genetics:  
A;Gene: GDB: LMO3; RBTN12  
A;Cross-references: GDB:127816; OMIM:180386  
A;Map position: 12p13-12p13  
P;16-69/Domain: LIM metal-binding repeat homology <LIM2>  
Query Match 27.8%; Score 5; DB 2; Length 72;

QY 14 CAGCN 18  
|||  
Db 16 CAGCN 20

Search completed: March 26, 2005, 17:25:59  
Job time : 18.6429 secs

```
Query Match      33.3%; Score 6; DB 1; Length 2483;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 15  
C93370  
rhombotin homolog 3 - human (fragment)  
C/Species: Homo sapiens (man)  
C/Date: 20-Mar-1992 #sequence\_revision 20-Mar-1992 #text\_change 15-Mar-2004  
C/Accession: C93370  
R/Boehm, T.; Foroni, L.; Kaneko, Y.; Perutz, M.F.; Rabbitts, T.H.

*This Page Blank (uspto)*

| Result No. | Query |       |        | DB | ID                   | Description       |
|------------|-------|-------|--------|----|----------------------|-------------------|
|            | Score | Match | Length |    |                      |                   |
| 1          | 166.5 | 30.4  | 115    | 2  | US-07-903-029-5      | Sequence 5, Appli |
| 2          | 163.5 | 29.8  | 115    | 2  | US-07-903-029-4      | Sequence 4, Appli |
| 3          | 135   | 24.6  | 115    | 2  | US-07-903-029-6      | Sequence 6, Appli |
| 4          | 81    | 14.8  | 15     | 1  | US-08-145-940-2      | Sequence 1, Appli |
| 5          | 81    | 14.8  | 16     | 1  | US-08-145-940-1      | Sequence 2, Appli |
| 6          | 81    | 14.8  | 16     | 2  | US-08-583-447A-56    | Sequence 56, Appl |
| 7          | 78    | 14.2  | 15     | 2  | US-08-583-447A-55    | Sequence 55, Appl |
| 8          | 69.5  | 12.7  | 1940   | 2  | US-08-844-271-30     | Sequence 30, Appl |
| 9          | 69.5  | 12.7  | 1940   | 4  | US-09-077-955-34     | Sequence 34, Appl |
| 10         | 68.5  | 12.5  | 221    | 3  | US-08-896-933-29     | Sequence 29, Appl |
| 11         | 68.5  | 12.5  | 221    | 3  | US-09-314-235-29     | Sequence 29, Appl |
| 12         | 68.5  | 12.5  | 221    | 4  | US-09-708-008B-29    | Sequence 29, Appl |
| 13         | 67.5  | 12.3  | 251    | 3  | US-09-144-776B-16    | Sequence 16, Appl |
| 14         | 67.5  | 12.3  | 251    | 4  | US-08-882-431B-16    | Sequence 16, Appl |
| 15         | 67    | 12.2  | 770    | 1  | US-08-525-654A-1     | Sequence 1, Appli |
| 16         | 67    | 12.2  | 771    | 1  | US-08-525-654A-3     | Sequence 3, Appli |
| 17         | 66.5  | 12.1  | 218    | 4  | US-09-893-737-28     | Sequence 28, Appl |
| 18         | 66.5  | 12.1  | 719    | 4  | US-09-949-016-7766   | Sequence 7766, Ap |
| 19         | 66    | 12.0  | 674    | 4  | US-09-328-352-6455   | Sequence 6455, Ap |
| 20         | 64.5  | 11.8  | 503    | 4  | US-09-252-991A-22790 | Sequence 22790, A |
| 21         | 64    | 11.7  | 211    | 4  | US-09-540-236-2321   | Sequence 2321, Ap |
| 22         | 64    | 11.7  | 700    | 4  | US-09-502-540-14225  | Sequence 14225, A |
| 23         | 64    | 11.7  | 917    | 2  | US-08-588-983-16     | Sequence 16, Appl |
| 24         | 64    | 11.7  | 917    | 2  | US-08-588-976-16     | Sequence 16, Appl |
| 25         | 63.5  | 11.6  | 498    | 1  | US-08-357-598-9      | Sequence 9, Appli |
| 26         | 63.5  | 11.6  | 498    | 2  | US-09-003-289-9      | Sequence 9, Appli |
| 27         | 63.5  | 11.6  | 498    | 5  | PCR-US95-16435-9     | Sequence 9, Appli |

```

Query Match      30.4%; Score 166.5; DB 2; Length 115;
Best Local Similarity 35.1%; Pred. No. 9.5e-13;
Matches 40; Conservative 15; Mismatches 32; Indels 27; Gaps 4;

QY 7 AAVALLLVLOSAGQVYIKYHGFQVQLSEYVKLNEELKEQMSDPQQQKSG----- 56
    ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
    15 AAAL-----AGGVTVQDGNFSLSEYVKLKDLOE-----PQEPVGVKLRNFATPG 61

DB 15 AAAL-----AGGVTVQDGNFSLSEYVKLKDLOE-----PQEPVGVKLRNFATPG 61

OV 57 -LLPDVCYNPALPDLQDPVCAQSEASTFKALRTATD--ECSLCINVACTGC 106

```





TELEFAX: (708)470-6881  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-145-940-2

Query Match 14.8%; Score 81; DB 1; Length 15;  
Best Local Similarity 85.7%; Pred. No. 0.002;  
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 93 DECELCINVACTGC 106  
Db 1 DDCELCNVACTGC 14

## RESULT 5

US-08-145-940-1  
Sequence 1, Application US/08145940  
Patent No. 5489670  
GENERAL INFORMATION:  
APPLICANT: Currie, Mark G.  
APPLICANT: Kita, Toshinori  
APPLICANT: Smith, Christine E.  
APPLICANT: Fok, Kam F.  
TITLE OF INVENTION: Human Uroguanylin  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,  
ADDRESSEE: Corporate Patent Dept.  
STREET: P. O. Box 5110  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60680  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/145,940  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Bennett, Dennis A.  
REGISTRATION NUMBER: 34,547  
REFERENCE/DOCKET NUMBER: 07-21(808)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (708)470-6501  
TELEFAX: (708)470-6881  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-145-940-1

Query Match 14.8%; Score 81; DB 1; Length 16;  
Best Local Similarity 85.7%; Pred. No. 0.0022;  
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 93 DECELCINVACTGC 106  
Db 2 DDCELCNVACTGC 15

## RESULT 6

US-08-583-447A-56  
Sequence 56, Application US/08583447A

Patent No. 5879656  
GENERAL INFORMATION:  
APPLICANT: Waldman, Scott A.  
TITLE OF INVENTION: ST Receptor Binding Compounds and  
METHODS OF USING THE SAME  
NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5879656ris  
STREET: One Liberty Place, 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: WordPerfect 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/583,447A  
FILING DATE: 05-JAN-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/141,892  
FILING DATE: 26-OCT-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: DeLuca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-1702  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 56:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-583-447A-56

Query Match 14.8%; Score 81; DB 2; Length 16;  
Best Local Similarity 85.7%; Pred. No. 0.0022;  
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 93 DECELCINVACTGC 106  
Db 2 DDCELCNVACTGC 15

## RESULT 7

US-08-583-447A-55  
Sequence 55, Application US/08583447A  
Patent No. 5879656  
GENERAL INFORMATION:  
APPLICANT: Waldman, Scott A.  
TITLE OF INVENTION: ST Receptor Binding Compounds and  
METHODS OF USING THE SAME  
NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5879656ris  
STREET: One Liberty Place, 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: WordPerfect 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/583,447A

```

; FILING DATE: 05-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/141,892
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1702
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-583-447A-55

Query Match 14.2%; Score 78; DB 2; Length 15;
Best Local Similarity 85.7%; Pred. No. 0.0048;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 93 DECELCINVACTGC 106
Db 2 EDCELCINVACTGC 15

RESULT 8
US-08-644-271-30
; Sequence 30, Application US/08644271
; Patent No. 5814478
; GENERAL INFORMATION:
; APPLICANT: Valenzuela, et al.
; TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS
; AND LIGANDS
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill Road
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/644,271
; FILING DATE: 10-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 60/008,657
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Cobert, Robert J
; REGISTRATION NUMBER: 36,108
; REFERENCE/DOCKET NUMBER: REG 195A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 914-345-7400
; TELEFAX: 914-345-7721
; TELEX:
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1940 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

; FEATURE:
; NAME/KEY: Rat Agrin
; LOCATION: 1...1940
; OTHER INFORMATION:
; US-08-644-271-30

Query Match 12.7%; Score 69.5; DB 2; Length 1940;
Best Local Similarity 26.2%; Pred. No. 39;
Matches 21; Conservative 7; Mismatches 19; Indels 33; Gaps 3;

Qy 49 DPQOKSGLLPDVCYNPALDLDPVCASQEAASFTFKALRTIATDECEL-----97
Db 79 EPSVEDPGRASCVCCKNACPATVAPCGSD--ASTY-----SNECELQRAQCNCQORRI 129

Qy 98 -----CINVACT 104
Db 130 RLLRQGPCGSRDPCANVTCS 149

RESULT 9
US-09-077-955-34
; Sequence 34, Application US/09077955A
; Patent No. 6413740
; GENERAL INFORMATION:
; APPLICANT: Valenzuela et al., David M.
; TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS AND LIGANDS
; FILE REFERENCE: REG195-B-PCT-US
; CURRENT APPLICATION NUMBER: US/09/077,955A
; CURRENT FILING DATE: 1998-09-10
; EARLIER APPLICATION NUMBER: PCT/US96/20696
; EARLIER FILING DATE: 1996-12-13
; EARLIER APPLICATION NUMBER: 08/644,271
; EARLIER FILING DATE: 1996-05-10
; EARLIER APPLICATION NUMBER: 60/008,657
; EARLIER FILING DATE: 1995-12-15
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 34
; LENGTH: 1940
; TYPE: PRT
; ORGANISM: Rattus sp.
; US-09-077-955-34

Query Match 12.7%; Score 69.5; DB 4; Length 1940;
Best Local Similarity 26.2%; Pred. No. 39;
Matches 21; Conservative 7; Mismatches 19; Indels 33; Gaps 3;

Qy 49 DPQOKSGLLPDVCYNPALDLDPVCASQEAASFTFKALRTIATDECEL-----97
Db 79 EPSVEDPGRASCVCCKNACPATVAPCGSD--ASTY-----SNECELQRAQCNCQORRI 129

Qy 98 -----CINVACT 104
Db 130 RLLRQGPCGSRDPCANVTCS 149

RESULT 10
US-08-896-933-29
; Sequence 29, Application US/08896933
; Patent No. 6221351
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; AND RELATED COMPOUNDS
; FILE REFERENCE: 09629/005002
; CURRENT APPLICATION NUMBER: US/08/896,933
; CURRENT FILING DATE: 1997-07-18
; EARLIER APPLICATION NUMBER: 08/252,978
; EARLIER FILING DATE: 1994-06-02
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 221
```

US-09-144-776B-16

Query Match 12.3%; Score 67.5; DB 3; Length 251;  
Best Local Similarity 23.6%; Pred. No. 4.3; Indels 35; Gaps 4;  
Matches 25; Conservative 16; Mismatches 30; Indels 35; Gaps 4;  
QY 15 VLSAQGVYIKYHGFQVLESVKLNELEBKQMSDPQQKSGLLPDVCYNPALPL--DLQ 72  
Db 44 LVKNLQNIYFLYEGDPVTHENVKSVQLRSH-----DLIYVSGPNYDKLK 89  
QY 73 PVCASQEAATFFKALRTIATDE-----CELCINVACTGC 106  
Db 90 TELKNQEMATLFK-----DKNVDIYGVYHYHLVCYLCAERSAC 128

RESULT 14  
US-08-882-431B-16  
; Sequence 16, Application US/08882431B  
; Patent No. 6713284  
; GENERAL INFORMATION:  
; APPLICANT: Robert G. Ulrich,  
; APPLICANT: Mark A. Olson  
; APPLICANT: Sina Bavari  
; TITLE OF INVENTION: Bacterial Superantigen  
; TITLE OF INVENTION: Vaccines  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Charles H. Harris  
; STREET: US Army MPMC -504 Scott Street MCMR-JA (Charles H. Harris-Patent Atty)  
; CITY: FORT DETRICK  
; STATE: MARYLAND  
; COUNTRY: USA  
; ZIP: 21702-5012  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Macintosh 7.5  
; SOFTWARE: Microsoft Word 6.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/882,431B  
; FILING DATE: June 25, 1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Charles H. Harris  
; REGISTRATION NUMBER: 34,616  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 619-2065  
; TELEFAX: (301) 619-7714  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 251  
; TYPE: Amino Acid  
; STRANDEDNESS: Unknown  
; TOPOLOGY: Unknown  
; MOLECULE TYPE: Peptide  
US-08-882-431B-16

Query Match 12.3%; Score 67.5; DB 4; Length 251;  
Best Local Similarity 23.6%; Pred. No. 4.3;  
Matches 25; Conservative 16; Mismatches 30; Indels 35; Gaps 4;  
QY 15 VLSAQGVYIKYHGFQVLESVKLNELEBKQMSDPQQKSGLLPDVCYNPALPL--DLQ 72  
Db 44 LVKNLQNIYFLYEGDPVTHENVKSVQLRSH-----DLIYVSGPNYDKLK 89  
QY 73 PVCASQEAATFFKALRTIATDE-----CELCINVACTGC 106  
Db 90 TELKNQEMATLFK-----DKNVDIYGVYHYHLVCYLCAERSAC 128

RESULT 15

US-08-525-654A-1  
; Sequence 1, Application US/08525654A  
; Patent No. 5736356  
; GENERAL INFORMATION:  
; APPLICANT: SANO, KOHICHIRO  
; APPLICANT: KUMAZAWA, YOSHIYUKI  
; APPLICANT: YASEUDA, HISASHI  
; APPLICANT: SEGURO, KATSUYA  
; APPLICANT: MOTOKI, MASAO  
; TITLE OF INVENTION: TRANSLUTAMINASE ORIGINATED FROM  
; TITLE OF INVENTION: CRASSOSTREA GIGAS  
; NUMBER OF SEQUENCES: 150  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/525,654A  
; FILING DATE: 28-SEP-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6/8283  
; FILING DATE: 28-JAN-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7/3876  
; FILING DATE: 13-JAN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 10-760-0 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 770 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; ORIGINAL SOURCE:  
; ORGANISM: Crassostrea gigas  
US-08-525-654A-1

Query Match 12.2%; Score 67; DB 1; Length 770;  
Best Local Similarity 23.7%; Pred. No. 22;  
Matches 22; Conservative 19; Mismatches 38; Indels 14; Gaps 3;  
QY 17 QSAQGVYIKYHGFQVLESVKLNE-----LEEKQMSDPQQKSGLLPDVCYNPALPLD 70  
Db 95 EKAGNIYRGQAFVIRVEFDREVNSDHDVILLOFTYGRPQESKGTVI-----RIPLD 147  
QY 71 LQFVCASQEAAT-FAKALRTIATDELCINVA 102  
Db 148 LKPTTKTSVDTETWFAEVKNIAKGLECAITSA 180

Search completed: March 26, 2005, 16:51:12  
Job time : 31.7226 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2005, 16:48:40 ; Search time 105.613 Seconds  
(without alignments)  
332.314 Million cell updates/sec

Title: US-10-775-481A-55

Perfect score: 548

Sequence: 1 MSGQLWAVVLLVLSQA.....LRTIATDECELCINVACTGC 106

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1407402 seqs, 331100923 residues

Total number of hits satisfying chosen parameters: 1407402

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*

2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*

3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*

4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*

5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*

6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*

7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*

8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*

9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*

10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*

11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*

12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*

13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*

14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*

15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*

16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*

17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*

18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*

19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*

20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description        |
|------------|-------|-------------|--------|----|--------------------|
| 1          | 356   | 65.0        | 112    | 17 | US-10-479-606-5    |
| 2          | 252.5 | 46.1        | 109    | 17 | US-10-479-606-6    |
| 3          | 166.5 | 30.4        | 115    | 9  | US-09-981-353-61   |
| 4          | 166.5 | 30.4        | 115    | 14 | US-10-235-994-22   |
| 5          | 166.5 | 30.4        | 115    | 14 | US-10-262-473-12   |
| 6          | 166.5 | 30.4        | 115    | 17 | US-10-479-606-4    |
| 7          | 120   | 21.9        | 102    | 14 | US-10-262-473-14   |
| 8          | 120   | 21.9        | 108    | 14 | US-10-262-473-16   |
| 9          | 97.5  | 17.8        | 87     | 9  | US-09-925-299-1162 |
| 10         | 97.5  | 17.8        | 87     | 10 | US-09-925-299-1162 |
| 11         | 84    | 15.3        | 16     | 14 | US-10-107-814-20   |
| 12         | 84    | 15.3        | 85     | 9  | US-09-925-299-1527 |
| 13         | 84    | 15.3        | 85     | 10 | US-09-925-299-1527 |

|    |      |      |      |    |                      |                   |
|----|------|------|------|----|----------------------|-------------------|
| 14 | 81   | 14.8 | 16   | 14 | US-10-107-814-1      | Sequence 1, Appli |
| 15 | 81   | 14.8 | 16   | 14 | US-10-197-954-141    | Sequence 141, App |
| 16 | 81   | 14.8 | 16   | 15 | US-10-621-684-56     | Sequence 56, Appl |
| 17 | 81   | 14.8 | 16   | 17 | US-10-479-606-2      | Sequence 2, Appli |
| 18 | 81   | 14.8 | 16   | 17 | US-10-760-085-141    | Sequence 141, App |
| 19 | 78   | 14.2 | 14   | 14 | US-10-107-814-21     | Sequence 21, Appl |
| 20 | 78   | 14.2 | 15   | 15 | US-10-621-684-55     | Sequence 55, Appl |
| 21 | 75   | 13.7 | 169  | 16 | US-10-437-963-156045 | Sequence 156045,  |
| 22 | 73.5 | 13.4 | 402  | 16 | US-10-437-963-132955 | Sequence 132955,  |
| 23 | 70.5 | 12.9 | 242  | 15 | US-10-424-599-275208 | Sequence 275208,  |
| 24 | 69.5 | 12.7 | 458  | 15 | US-10-369-493-17284  | Sequence 17284, A |
| 25 | 69.5 | 12.7 | 742  | 16 | US-10-437-963-185159 | Sequence 185159,  |
| 26 | 69.5 | 12.7 | 1937 | 15 | US-10-072-012-792    | Sequence 792, App |
| 27 | 69.5 | 12.7 | 1940 | 13 | US-10-016-283-34     | Sequence 34, Appl |
| 28 | 69.5 | 12.7 | 1940 | 15 | US-10-072-012-791    | Sequence 791, App |
| 29 | 69.5 | 12.7 | 1959 | 15 | US-10-072-012-790    | Sequence 790, App |
| 30 | 69   | 12.6 | 15   | 17 | US-10-479-606-3      | Sequence 3, Appli |
| 31 | 69   | 12.6 | 866  | 16 | US-10-437-963-183310 | Sequence 183310,  |
| 32 | 68.5 | 12.5 | 358  | 15 | US-10-335-977-6075   | Sequence 6075, Ap |
| 33 | 68.5 | 12.5 | 431  | 15 | US-10-276-774-1563   | Sequence 1563, Ap |
| 34 | 68.5 | 12.5 | 547  | 8  | US-08-834-705-2      | Sequence 2, Appli |
| 35 | 68.5 | 12.5 | 547  | 15 | US-10-282-122A-59002 | Sequence 59002, A |
| 36 | 68.5 | 12.5 | 549  | 15 | US-10-335-977-6076   | Sequence 6076, Ap |
| 37 | 68.5 | 12.5 | 549  | 15 | US-10-335-977-6077   | Sequence 6077, Ap |
| 38 | 68   | 12.4 | 72   | 17 | US-10-796-719-21     | Sequence 21, Appl |
| 39 | 68   | 12.4 | 220  | 14 | US-10-002-784A-26    | Sequence 26, Appl |
| 40 | 68   | 12.4 | 235  | 16 | US-10-437-963-203716 | Sequence 203716,  |
| 41 | 68   | 12.4 | 468  | 14 | US-10-002-784A-27    | Sequence 27, Appl |
| 42 | 68   | 12.4 | 900  | 16 | US-10-437-963-105276 | Sequence 105276,  |
| 43 | 67.5 | 12.3 | 251  | 8  | US-08-882-431-16     | Sequence 16, Appl |
| 44 | 67.5 | 12.3 | 251  | 14 | US-10-002-784A-16    | Sequence 16, Appl |
| 45 | 67.5 | 12.3 | 1074 | 15 | US-10-389-566-2371   | Sequence 2371, Ap |

ALIGNMENTS

RESULT 1

US-10-479-606-5

Sequence 5, Application US/10479606

Publication No. US20050032684A1

GENERAL INFORMATION:

APPLICANT: Cetin, Yalcin

APPLICANT: Savas, Yuksele

TITLE OF INVENTION: Guanlylate-cyclase C ligand, administered via the airways, for the treatment of respiratory airway problems

FILE REFERENCE: 03100192aa

CURRENT APPLICATION NUMBER: US/10/479,606

CURRENT FILING DATE: 2003-12-04

PRIOR APPLICATION NUMBER: DE10127119.0

PRIOR FILING DATE: 2001-06-05

PRIOR APPLICATION NUMBER: PCT/DE02/02040

PRIOR FILING DATE: 2002-06-05

NUMBER OF SEQ ID NOS: 7

SOFTWARE: Patentin version 3.2

SEQ ID NO 5

LENGTH: 112

TYPE: PRT

ORGANISM: homo sapiens

US-10-479-606-5

Query Match 65.0%; Score 356; DB 17; Length 112;  
Best Local Similarity 65.7%; Pred. No. 1.5e+33;  
Matches 69; Conservative 14; Mismatches 22; Indels 0; Gaps 0;

Qy 2 SGSQLWAVVLLVLSQAQGVYKYHGFQVLESVKKLENEEKQMSDPQOKSGLLPDV 61

Db 7 SGLPGVAVVLLVLSQTSQVYIQGFRVQLESKMLSLDLAQWAPSPRLQASLLPAV 66

Qy 62 CYNPALPDLQPCVCAQEAATFKALRTIATDECELCINVACTGC 106

Db 67 CHHPALPDLQPCVCAQEAATFKALRTIATDECELCINVACTGC 111



```
Query Match 30.4%; Score 166.5; DB 14; Length 115;
Best Local Similarity 35.1%; Pred. No. 2.3e-11;
Matches 40; Conservative 15; Mismatches 32; Indels 27; Gaps 4;

QY 7 WAAVLLLVLSAQGVYIKYHGQVQLSESVKLNLEEKQMSDPQOQKSG-----56
Db 15 WAAL-----AGGVTVQDGNFSLSESVKLNKDLQE-----POEPRVGKLRNFAPIPG 61

QY 57 --LLPDVCYNPALPLDLQPVCSQAASSTFKALRTIATD--ECELINVACTGC 106
Db 62 BPVVPILCSNPFPPEELPKCKEPNAQEILQRLLEIAEDPGTCEICAVAACTGC 115

RESULT 6
US-10-479-606-4
; Sequence 4, Application US/10479606
; Publication No. US20050032684A1
; GENERAL INFORMATION:
; APPLICANT: Cetin, Yalcin
; APPLICANT: Savas, Yukeel
; TITLE OF INVENTION: Guanylate-cyclase C ligand, administered via the airways, for the
; FILE REFERENCE: 03100192aa
; CURRENT APPLICATION NUMBER: US/10/479,606
; CURRENT FILING DATE: 2003-12-04
; PRIOR APPLICATION NUMBER: DE10127119.0
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/DE02/02040
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 4
; LENGTH: 115
; TYPE: PRT
; ORGANISM: rat or homo sapiens (guanylin)
; US-10-479-606-4

Query Match 30.4%; Score 166.5; DB 17; Length 115;
Best Local Similarity 35.1%; Pred. No. 2.3e-11;
Matches 40; Conservative 15; Mismatches 32; Indels 27; Gaps 4;

QY 7 WAAVLLLVLSAQGVYIKYHGQVQLSESVKLNLEEKQMSDPQOQKSG-----56
Db 15 WAAL-----AGGVTVQDGNFSLSESVKLNKDLQE-----POEPRVGKLRNFAPIPG 61

QY 57 --LLPDVCYNPALPLDLQPVCSQAASSTFKALRTIATD--ECELINVACTGC 106
Db 62 BPVVPILCSNPFPPEELPKCKEPNAQEILQRLLEIAEDPGTCEICAVAACTGC 115

RESULT 7
US-10-262-473-14
; Sequence 14, Application US/10262473
; Publication No. US20030199442A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook, John,
; APPLICANT: Burgess, Catherine,
; APPLICANT: Gorman, Linda,
; APPLICANT: Guo, Xiaojia,
; APPLICANT: Lepley, Denise,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rastelli, Luca,
; APPLICANT: Reiger, Daniel,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-462B
; CURRENT APPLICATION NUMBER: US/10/262,473
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/349,575
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 16
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-262-473-16

Query Match 21.9%; Score 120; DB 14; Length 108;
Best Local Similarity 33.0%; Pred. No. 5.9e-06;
Matches 34; Conservative 12; Mismatches 39; Indels 18; Gaps 5;

QY 7 WAAVLLLVLSAQGVYIKYHGQVQLSESVKLNLEEKQMSDPQOQKSGLLPDVCYNP 65
Db 18 WAAL-----AGGVTVQDGNFSLSESVKLNKDLQE-----VVPILCSNP 62

QY 66 ALPLDLQPVCSQAASSTFKALRTIATD--ECELINVACTGC 106
```

```
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/349,575
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 14
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-262-473-14

Query Match 21.9%; Score 120; DB 14; Length 102;
Best Local Similarity 33.0%; Pred. No. 5.9e-06;
Matches 34; Conservative 12; Mismatches 39; Indels 18; Gaps 5;

QY 7 WAAVLLLVLSAQGVYIKYHGQVQLSESVKLNLEEKQMSDPQOQKSGLLPDVCYNP 65
Db 15 WAAL-----AGGVTVQDGNFSLSESVKLNKDLQE-----VVPILCSNP 59

QY 66 ALPLDLQPVCSQAASSTFKALRTIATD--ECELINVACTGC 106
Db 60 NFPEELPKCKEPNAQEILQRLLEIAEDPGTCEICAVAACTGC 102
```

```
RESULT 8
US-10-262-473-16
; Sequence 16, Application US/10262473
; Publication No. US20030199442A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook, John,
; APPLICANT: Burgess, Catherine,
; APPLICANT: Gorman, Linda,
; APPLICANT: Guo, Xiaojia,
; APPLICANT: Lepley, Denise,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rastelli, Luca,
; APPLICANT: Reiger, Daniel,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-462B
; CURRENT APPLICATION NUMBER: US/10/262,473
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/349,575
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 16
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-262-473-16
```

```
Query Match 21.9%; Score 120; DB 14; Length 108;
Best Local Similarity 33.0%; Pred. No. 5.9e-06;
Matches 34; Conservative 12; Mismatches 39; Indels 18; Gaps 5;

QY 7 WAAVLLLVLSAQGVYIKYHGQVQLSESVKLNLEEKQMSDPQOQKSGLLPDVCYNP 65
Db 18 WAAL-----AGGVTVQDGNFSLSESVKLNKDLQE-----VVPILCSNP 62

QY 66 ALPLDLQPVCSQAASSTFKALRTIATD--ECELINVACTGC 106
```





;; CURRENT FILING DATE: 2001-08-10  
;; PRIOR APPLICATION NUMBER: PCT/US00/05883  
;; PRIOR FILING DATE: 2000-03-08  
;; PRIOR APPLICATION NUMBER: 60/124,270  
;; PRIOR FILING DATE: 1999-03-12  
;; NUMBER OF SEQ ID NOS: 1556  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 1527  
;; LENGTH: 85  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: SITE  
;; LOCATION: (1)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (3)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (9)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (28)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (31)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (32)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (34)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (36)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (38)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (39)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (42)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (43)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (47)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (52)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (53)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (55)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (60)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (66)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (70)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE

;; LOCATION: (71)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (83)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-299-1527  
Query Match 15.3%; Score 84; DB 9; Length 85;  
Best Local Similarity 33.8%; Pred. NO. 0.07;  
Matches 23; Conservative 10; Mismatches 35; Indels 0; Gaps 0;  
QY 2 SCSQLWAARVILLVLLVLSQAGVYIKYHGFVQVLESVKKLNELEEKQMSDPQOKSGLLPDV 61  
Db 17 SGLTGLAEVLXLLXLTQXGXQYXXFRVXLESXXLLDLEAXWAPSPXLEAXXLLAAV 76  
QY 62 CYNPALPL 69  
Db 77 CHHPALXL 84  
RESULT 13  
US-09-925-299-1527  
;; Sequence 1527, Application US/09925299  
;; Publication No. US20030040617A9  
;; GENERAL INFORMATION:  
;; APPLICANT: Rosen et al.  
;; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
;; FILE REFERENCE: PA102  
;; CURRENT APPLICATION NUMBER: US/09/925,299  
;; CURRENT FILING DATE: 2001-08-10  
;; PRIOR APPLICATION NUMBER: PCT/US00/05883  
;; PRIOR FILING DATE: 2000-03-08  
;; PRIOR APPLICATION NUMBER: 60/124,270  
;; PRIOR FILING DATE: 1999-03-12  
;; NUMBER OF SEQ ID NOS: 1556  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 1527  
;; LENGTH: 85  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: SITE  
;; LOCATION: (1)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (3)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (9)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (28)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (29)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (31)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (32)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (34)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (36)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (38)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
;; NAME/KEY: SITE  
;; LOCATION: (39)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (42)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (43)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (47)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (52)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (53)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (55)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (60)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (66)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (70)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (71)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (83)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-299-1527

Query Match 15.3%; Score 84; DB 10; Length 85;  
Best Local Similarity 33.8%; Pred. No. 0.07;  
Matches 23; Conservative 10; Mismatches 35; Indels 0; Gaps 0;  
Qy 2 SGSQLWAVLLLVLSAQGVYIKYGFQVLESVKLNELEKQMSDPQQKSGLLPDV 61  
Db 17 SGLLTGLAEVLLXXLTQXGXQYXXFRVXLESXXLXDLAXWAPSPXLEAXXLLAAV 76  
Qy 62 CYNPALPL 69  
Db 77 CHHPALXL 84

RESULT 14  
US-10-107-814-1  
Sequence 1, Application US/10107814  
Publication No. US20030073628A1  
GENERAL INFORMATION:  
APPLICANT: SHAILUBHAI, KUNWAR  
APPLICANT: NIKIFOROVICH, GREGORY  
APPLICANT: JACOB, GARY S.  
TITLE OF INVENTION: GUANYLATE CYCLASE RECEPTOR AGONISTS FOR THE TREATMENT  
OF TISSUE INFLAMMATION AND CARCINOGENESIS  
FILE REFERENCE: 81361/284943/MAS  
CURRENT APPLICATION NUMBER: US/10/107,814  
CURRENT FILING DATE: 2002-03-28  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: DISULFID  
LOCATION: (4)..(12)  
NAME/KEY: DISULFID  
LOCATION: (7)..(15)

US-10-107-814-1

Query Match 14.8%; Score 81; DB 14; Length 16;  
Best Local Similarity 85.7%; Pred. No. 0.018;  
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 93 DECELCINVACTGC 106  
Db 2 DDCELCVNVACTGC 15

RESULT 15

US-10-197-954-141  
Sequence 141, Application US/10197954  
Publication No. US20030119021A1  
GENERAL INFORMATION:  
APPLICANT: K\*ster, Hubert  
APPLICANT: Siddiqi, Suhail  
APPLICANT: Little, Daniel  
TITLE OF INVENTION: Capture Compounds, Collections Thereof  
And Methods For Analyzing The Proteome And Complex  
Compositions  
TITLE OF INVENTION: Compositions  
FILE REFERENCE: 24743-2305  
CURRENT APPLICATION NUMBER: US/10/197,954  
CURRENT FILING DATE: 2002-07-16  
PRIOR APPLICATION NUMBER: 60/306,019  
PRIOR FILING DATE: 2001-07-16  
PRIOR APPLICATION NUMBER: 60/314,123  
PRIOR FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: 60/363,433  
PRIOR FILING DATE: 2002-03-11  
NUMBER OF SEQ ID NOS: 149  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 141  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-197-954-141

Query Match 14.8%; Score 81; DB 14; Length 16;  
Best Local Similarity 85.7%; Pred. No. 0.018;  
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 93 DECELCINVACTGC 106  
Db 2 DDCELCVNVACTGC 15

Search completed: March 26, 2005, 17:16:38  
Job time : 106.613 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2005, 16:37:49 ; Search time 22.438 Seconds  
(without alignments)  
454.541 Million cell updates/sec

Title: US-10-775-481A-55  
Perfect score: 548  
Sequence: 1 MSGSQLMAVLLLVLSQAQ.....LRTIATDELCINVACTGC 106

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description           |
|------------|-------|-------------|--------|----------|-----------------------|
| 1          | 356   | 65.0        | 112    | JC4651   | uroguanylin precursor |
| 2          | 167.5 | 30.6        | 116    | JC7620   | guanylin precursor    |
| 3          | 166.5 | 30.4        | 115    | A46279   | guanylin precursor    |
| 4          | 135   | 24.6        | 115    | JN0318   | guanylin precursor    |
| 5          | 134.5 | 24.5        | 116    | B46279   | alkanal monooxygen    |
| 6          | 74    | 13.5        | 354    | JQ0413   | coatamer protein z    |
| 7          | 73    | 13.3        | 189    | S52521   | proteinase inhibitor  |
| 8          | 72.5  | 13.2        | 228    | S45677   | heat-stable entero    |
| 9          | 70.5  | 12.9        | 72     | Q4EC1B   | response regulator    |
| 10         | 70.5  | 12.9        | 368    | A82249   | exotoxin type A pr    |
| 11         | 69.5  | 12.7        | 236    | S18786   | NADP-specific glut    |
| 12         | 69.5  | 12.7        | 458    | B83912   | c-fps proto oncoge    |
| 13         | 69.5  | 12.7        | 824    | I50618   | agrin - rat           |
| 14         | 69.5  | 12.7        | 1959   | 1 AGRT   | exotoxin type A pr    |
| 15         | 68.5  | 12.5        | 236    | S18783   | hypothetical prote    |
| 16         | 68.5  | 12.5        | 251    | S29659   | tetrahydropteroylt    |
| 17         | 68.5  | 12.5        | 342    | AC3027   | probable 60K inner    |
| 18         | 68.5  | 12.5        | 342    | G98257   | probable inner mem    |
| 19         | 68.5  | 12.5        | 547    | B64701   | heat-stable entero    |
| 20         | 68.5  | 12.5        | 549    | E71818   | protein-tyrosine k    |
| 21         | 68    | 12.4        | 72     | 1 QHEC4  | hypothetical prote    |
| 22         | 68    | 12.4        | 533    | 1 TVFVP  | ecdysone-induced p    |
| 23         | 67.5  | 12.3        | 1074   | 2 T01906 | phospholipase C (E    |
| 24         | 67    | 12.2        | 873    | 2 B53225 | hypothetical prote    |
| 25         | 67    | 12.2        | 1210   | 2 A48001 | lamina associated     |
| 26         | 66.5  | 12.1        | 357    | 2 B84678 | microfibril-associ    |
| 27         | 66.5  | 12.1        | 462    | 2 I61730 |                       |
| 28         | 66.5  | 12.1        | 873    | 1 TVFVS  |                       |
| 29         | 66    | 12.0        | 183    | 2 A49313 |                       |

|    |      |      |      |   |        |                    |
|----|------|------|------|---|--------|--------------------|
| 30 | 66   | 12.0 | 406  | 2 | T23277 | hypothetical prote |
| 31 | 66   | 12.0 | 1259 | 2 | T32901 | hypothetical prote |
| 32 | 65   | 11.9 | 235  | 2 | G72501 | phosphonates trans |
| 33 | 65   | 11.9 | 247  | 2 | B83708 | probable ATP-bindi |
| 34 | 65   | 11.9 | 274  | 2 | E91185 | hypothetical prote |
| 35 | 65   | 11.9 | 274  | 2 | C86032 | probable ATP-bindi |
| 36 | 65   | 11.9 | 274  | 2 | S47791 | hypothetical prote |
| 37 | 65   | 11.9 | 496  | 2 | D83614 | conserved hypothet |
| 38 | 64.5 | 11.8 | 495  | 2 | S28167 | leucine aminopepti |
| 39 | 64.5 | 11.8 | 711  | 1 | S28391 | protein-tyrosine-p |
| 40 | 64.5 | 11.8 | 711  | 2 | T45160 | hypothetical prote |
| 41 | 64   | 11.7 | 445  | 2 | T72251 | carbon-monoxide de |
| 42 | 64   | 11.7 | 799  | 2 | H70832 | hypothetical prote |
| 43 | 64   | 11.7 | 852  | 2 | T46091 | hexokinase (EC 2.7 |
| 44 | 64   | 11.7 | 917  | 1 | S15885 | ATP-dependent RNA  |
| 45 | 64   | 11.7 | 993  | 2 | E71392 |                    |

ALIGNMENTS

RESULT 1

JC4651

uroguanylin precursor - human

N:Alternate names: Guanylyl cyclase activating peptide II

C:Species: Homo sapiens (man)

C>Date: 10-May-1996 #sequence revision 19-Jul-1996 #text\_change 09-Jul-2004

C:Accession: JC4651; S63702; S68052

R:Miyazato, M.; Nakazato, M.; Yamaguchi, H.; Date, Y.; Kojima, M.; Kangawa, K.; Matsuo,

Biochem. Biophys. Res. Commun. 219, 644-648, 1996

A:Title: Cloning and characterization of a cDNA encoding a precursor for human uroguanylin

A:Reference number: JC4651; MUID:96193705; PMID:8605041

A:Accession: JC4651

A:Molecule type: mRNA

A:Residues: 1-112 <M>

A:Cross-references: UNIPROT:Q16661; GB:U34279; NID:gl236798; PIDN:AAC50416.1; PID:gl2367

R: Hill, O.; Cetin, Y.; Cieslak, A.; Maegert, H.J.; Forssmann, W.G.

Biochim. Biophys. Acta 1253, 146-149, 1995

A:Title: A new human guanylate cyclase-activating peptide (GCAP-II, uroguanylin): precu

A:Reference number: S63702; MUID:96106424; PMID:8519795

A:Accession: S63702

A:Molecule type: mRNA

A:Residues: 1-112 <HML>

A:Cross-references: EMBL:Z50753; NID:9974823; PIDN:CAA90629.1; PID:g974824

R: Hess, R.; Kuhn, M.; Schulz-Knappe, P.; Raida, M.; Fuchs, M.; Klodt, J.; Adermann, K.;

FEBS Lett. 374, 34-38, 1995

A:Title: GCAP-II: isolation and characterization of the circulating form of human urogu

A:Reference number: S68052; MUID:96049550; PMID:7589507

A:Accession: S68052

A:Molecule type: protein

A:Residues: 89-99,'X',101-102,'X',104-107,'X',109-110,'X',112 <HES>

C:Comment: This protein, a member of the guanylin peptide family, is an endogenous activ

C:Superfamily: guanylin

C:Keywords: intestine

F:1-26/Domain: signal sequence #status predicted <SIG>

F:27-112/Product: uroguanylin #status predicted <MAT>

Query Match 65.0%; Score 356; DB 2; Length 112;

Best Local Similarity 65.7%; Pred. No. 9.8e-30;

Matches 69; Conservative 14; Mismatches 22; Indels 0; Gaps 0;

Qy 2 SGSQLMAVLLLVLSQAQGVYIKYHGFVQLSEKVKLNLEEKQMSDPQOKSGLLPDV 61

Db 7 SGLPGVAVLLLVLLQTSVYIQGFRVQLSEKVKLNLEEKQMSDPQOKSGLLPDV 66

Qy 62 CYNPALPLDLPVCASQEAATFKALRTIATDELCINVACTGC 106

Db 67 CHHPALPDLPVCASQEAATFKALRTIATDELCINVACTGC 111

RESULT 2

JC7620  
guanylin precursor, long form - European eel  
C;Species: Anguilla anguilla (European eel)  
C;Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 03-Aug-2001  
R;Comrie, M.M.; Cutler, C.P.; Cramb, G.  
Biochem. Biophys. Res. Commun. 281, 1078-1085, 2001  
A;Title: Cloning and expression of guanylin from the European eel (*Anguilla anguilla*).  
A;Reference number: JC7620; MUID:21139737; PMID:11243845  
A;Accession: JC7620  
A;Molecule type: mRNA  
A;Residues: 1-116 <COM>  
A;Cross-references: GB:A301673  
C;Comment: This protein, a member of a family of heat-stable peptides, is a potent extra-  
axis. This peptide signalling system plays a role in osmoregulation in euryhaline teleost  
C;Superfamily: guanylin  
C;Keywords: heat-stable protein; osmoregulation  
F;1-28/Domain: signal sequence #status predicted <SIG>  
F;29-116/Product: guanylin precursor, long form #status predicted <MAT>  
F;33-39/Region: homologous #status predicted  
F;69-114/Region: highly conserved #status predicted

Query Match 30.6%; Score 167.5; DB 2; Length 116;  
Best Local Similarity 41.9%; Pred. No. 3.6e-10;  
Matches 44; Conservative 10; Mismatches 44; Indels 7; Gaps 4;

QY 8 AAVLLILV---LQSAQGVYIKYHGFQVLESVKLINE-LEEKOMSDPQOKSGLLPDVCY 63  
Db 12 AVLLILAFCLQESQGVYMDGDLSPLEAVKVLKLGANTMSTPHPNLG-SHAVCS 70

QY 64 NPALPLDLPQVCASQEAATFFKALRTIAT--DECELCINVACTGC 106  
Db 71 NPHLPAEFLPVCREGASALFNRLVDIITPDPDCEICANNACTGC 115

RESULT 3  
A46279  
guanylin precursor [validated] - human  
C;Species: Homo sapiens (man)  
C;Date: 21-Sep-1993 #sequence\_revision 26-May-1995 #text\_change 09-Jul-2004  
C;Accession: A46279; S29228; S29807  
R;de Sauvage, F.J.; Keshav, S.; Kuang, W.J.; Gillett, N.; Henzel, W.; Goeddel, D.V.  
Proc. Natl. Acad. Sci. U.S.A. 89, 9089-9093, 1992  
A;Title: Precursor structure, expression, and tissue distribution of human guanylin.  
A;Reference number: A46279; MUID:93028409; PMID:1409606  
A;Accession: A46279  
A;Molecule type: mRNA  
A;Residues: 1-115 <DE1>  
A;Cross-references: UNIPROT:Q02747; GB:M95174; NID:G306823; PIDN:AAA58625.1; PID:G306824  
A;Note: sequence extracted from NCBI backbone (NCBIN:115377, NCBIPI:115378)  
R;Wiegand, R.C.; Kato, J.; Huang, M.D.; Fok, K.F.; Kachur, J.F.; Currie, M.G.  
FEBS Lett. 311, 150-154, 1992  
A;Title: Human guanylin: cDNA isolation, structure, and activity.  
A;Reference number: S29228; MUID:93011964; PMID:1327879  
A;Accession: S29228  
A;Molecule type: mRNA  
A;Residues: 1-115 <WIE>  
A;Cross-references: GB:M97496; NID:G183414; PIDN:AAA35915.1; PID:G183415  
R;Kuhn, M.; Raida, M.; Adermann, K.; Schulz-Knappe, P.; Gerzer, R.; Heim, J.M.; Forssman  
FEBS Lett. 318, 205-209, 1993  
A;Title: The circulating bioactive form of human guanylin is a high molecular weight pep  
A;Reference number: S29807; MUID:93178628; PMID:8095028  
A;Accession: S29807  
A;Molecule type: protein  
A;Residues: 22-68 <KUH>  
A;Experimental source: plasma  
C;Comment: amino-terminal sequencing of mature form and molecular weight of mature form by  
n of the same receptor.  
C;Genetics:  
A;Gene: GDB:GUCA2  
A;Cross-references: GDB:136460; OMIM:139392  
A;Map position: lp35-1p34

C;Superfamily: guanylin  
C;Keywords: hormone; intestine  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-115/Product: guanylin #status experimental <MAT>

Query Match 30.4%; Score 166.5; DB 1; Length 115;  
Best Local Similarity 35.1%; Pred. No. 4.5e-10;  
Matches 40; Conservative 15; Mismatches 32; Indels 27; Gaps 4;

QY 7 WAAVLLLLVLSQAQGVYIKYHGFQVLESVKLINELEEKOMSDPQOKSG----- 56  
Db 15 WAAL-----AGVTVDQGNFSLSVKXKLDIQE-----PQPRVKLRNFAPIG 61

QY 57 --LLPDVCYNPALPLDLPQVCASQEAATFFKALRTIATD--ECELINVACTGC 106  
Db 62 EPVVPILCSNPFPBELKELCKEPNAQEILQRLLEEIAEDPGTCEICAYAACTGC 115

RESULT 4  
JN0318  
guanylin precursor - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 04-Dec-1992 #sequence\_revision 26-May-1995 #text\_change 09-Jul-2004  
C;Accession: JN0318; A43345; A38184; S25489  
R;Wiegand, R.C.; Kato, J.; Currie, M.G.  
Biochem. Biophys. Res. Commun. 185, 812-817, 1992  
A;Title: Rat guanylin cDNA: characterization of the precursor of an endogenous activator  
A;Reference number: JN0318; MUID:92328783; PMID:1378267  
A;Accession: JN0318  
A;Molecule type: mRNA  
A;Residues: 1-115 <WIE>  
A;Cross-references: UNIPROT:P28902; GB:M93005; NID:G204540; PIDN:AAA41300.1; PID:G204541  
R;Schulz, S.; Christman, T.D.; Garbers, D.L.  
J. Biol. Chem. 267, 16019-16021, 1992  
A;Title: Cloning and expression of guanylin. Its existence in various mammalian tissues.  
A;Reference number: A43345; MUID:92355545; PMID:1379587  
A;Accession: A43345  
A;Molecule type: mRNA  
A;Residues: 1-115 <SCH>  
A;Cross-references: GB:M95493; NID:G204542; PIDN:AAA41302.1; PID:G204543  
A;Experimental source: intestine  
A;Note: sequence extracted from NCBI backbone (NCBIN:110474, NCBIPI:110476)  
R;Currie, M.G.; Fok, K.F.; Kato, J.; Moore, R.J.; Hamra, F.K.; Duffin, K.L.; Smith, C.E.  
Proc. Natl. Acad. Sci. U.S.A. 89, 947-951, 1992  
A;Title: Guanylin: an endogenous activator of intestinal guanylate cyclase.  
A;Reference number: A38184; MUID:92141235; PMID:1346555  
A;Accession: A38184  
A;Molecule type: protein  
A;Residues: 101-115 <CUR>  
A;Experimental source: jejunum  
A;Note: sequence extracted from NCBI backbone (NCBIPI:79480)  
A;Note: purification included boiling in acetic acid; peptide has activity but may repre  
R;Maegert, H.J.; Khun, M.; Kruhoffer, M.; Forssmann, W.G.  
submitted to the EMBL Data Library, August 1992  
A;Reference number: S25489  
A;Accession: S25489  
A;Molecule type: mRNA  
A;Residues: 101-115 <MAE>  
A;Cross-references: EMBL:X67669; NID:G56343; PIDN:CAA47901.1; PID:G56344  
C;Comment: Guanylin is an endogenous ligand for an intestine-specific receptor guanylyl  
n of the same receptor.  
C;Superfamily: guanylin  
C;Keywords: hormone; intestine  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-115/Product: guanylin #status predicted <MAT>

Query Match 24.6%; Score 135; DB 1; Length 115;  
Best Local Similarity 33.0%; Pred. No. 8.3e-07;  
Matches 37; Conservative 12; Mismatches 51; Indels 12; Gaps 3;

QY 7 WAAVLLLLVLSQAQGVYIKYHGFQVLESVKLINELEEKOMSDPQOKSGLL----- 58  
Db 4 WLLSVLCLLGLALVILVEGTVQDGLSPFLSVKQLKHLREVQEPETLMSHKKFAFLRLPKP 63

**QY** 59 --PDVYNPALPDLQPVCASQEAASTFKALRTIATD--ECELCINVACTGC 106  
::|::|::|::|:  
**Db** 64 VAPELCSQSAFPELRPLCEKPNABEILQRLEAIAODPNTCEICAYAACTGC 115  
|||  
  
**RESULT 5**

B46279  
guanylin precursor - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 22-Sep-1993 #sequence\_revision 26-May-1995 #text\_change 09-Jul-2004  
R:Accession: A55643; B46279  
R:Sciaky, D.; Koshiba, J.L.; Cohen, M.B.  
Genomics 24, 583-587, 1994  
A>Title: Genomic sequence of the murine guanylin gene.  
A:Reference number: A55643; PMID:95229161; PMID:7713512  
A:Accession: A55643  
A:Molecule type: DNA  
A:Residues: 1-116 <UNI>  
A:Cross-references: UNIPROT:P33680; GB:U60528; GB:U09741; NID:g1480667; PIDN:AABO5758.1; R:de Sauvage, F.J.; Keshav, S.; Kuang, W.J.; Gillett, N.; Henzel, W.; Goeddel, D.V. Proc. Natl. Acad. Sci. U.S.A. 89, 9089-9093, 1992  
A>Title: Precursor structure, expression, and tissue distribution of human guanylin.  
A:Reference number: A46279; PMID:93028409; PMID:1409606  
A:Accession: B46279  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-116 <DEI>  
A:Cross-references: GB:M95175; GVV:M9309282; PIDN:AAA37758.1; PID:g309283  
A>Note: sequence extracted from NCBI backbone (NCBIP:115379)  
C:Comment: Guanylin is an endogenous ligand for an intestine-specific receptor guanylyl n of the same receptor.  
C:Genetics:  
A:Introns: 25/3; 96/1  
C:Superfamily: guanylin  
C:Keywords: hormone; intestine  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-116/Product: guanylin #status predicted <MAT>

**Query Match** 24.5%; Score 134.5; DB 1; Length 116;  
**Best Local Similarity** 36.0%; Pred. No. 9.5e-07;  
**Matches** 41; Conservative 15; Mismatches 43; Indels 15; Gaps 5;

**QY** 8 AAVLLLVQLQA-----QGVIKYHGFQVOLESVKK---LNELEEKKMSDPQQKSGLL- 58  
::|::|::|::|:  
**Db** 3 ACVLVSCLLGALLVGVTVQGDLSFPLESVKKLKGRLREVQSPLRVSHKKFAPELLQ 62  
|||  
  
**QY** 59 ---PDVC-YNPALPDLQPVCASQEAASTFKALRTIATD--ECELCINVACTGC 106  
::|::|::|::|:  
**Db** 63 PVAPOLCSSHALPEARLPCKEPNAEBEILQRLEAIAODPNTCEICAYAACTGC 116  
|||  
  
**RESULT 6**

JQ0413  
alkanal monooxygenase (FMN-linked) (EC 1.14.14.3) alpha chain - Photobacterium leiognath  
N:Alternate names: bacterial luciferase; luciferase alpha chain  
C:Species: Photobacterium leiognathi  
C>Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 09-Jul-2004  
C:Accession: JQ0413; S06369; S01697  
R:Illarionov, B.A.; Blinov, V.M.; Donchenko, A.P.; Protopopova, M.V.; Karginov, V.A.; Me Gene 86, 89-94, 1990  
A>Title: Isolation of bioluminescent functions from Photobacterium leiognathi: analysis  
A:Reference number: PQ0049; PMID:90185252; PMID:2311938  
A:Accession: JQ0413  
A:Molecule type: DNA  
A:Residues: 1-354 <ILL>  
A:Cross-references: UNIPROT:P09140; GB:X080036; GB:M34564; GB:M38298; NID:g45589; PIDN:CA R:Illarionov, B.A.; Protopopova, M.V.; Karginov, V.A.; Mertvetsov, N.P.; Gitelson, I.I. Bioorg. Khim. 14, 412-415, 1988  
A>Title: Nucleotide sequence of genes of the luciferase alpha and beta subunits from Pho A:Reference number: S06369; PMID:88251495; PMID:13382442  
A:Accession: S06369  
A:Molecule type: DNA

A:Molecule type: protein  
A:Residues: 54-72 <AIM>  
R:Guzman-Verdusco, L.M.; Kupersztosch, Y.M.  
Infect. Immun. 57, 645-648, 1989  
A:Title: Rectification of two Escherichia coli heat-stable enterotoxin allele sequences  
A:Reference number: A30567; MUID:89108616; PMID:2643580  
A:Accession: A30567  
A:Molecule type: DNA  
A:Residues: 1-18,'A','20-24','AG','27-41','V','43-44','N','46','E','48','S','50-72 '<GUZ>  
A:Cross-references: GB:M18345; NID:g145862; PIDN:AAA23729.1; PID:g145863  
C:Comment: This is one of the type I heat-stable enterotoxins that are methanol-soluble.  
C:Genetics:  
A:Gene: st  
C:Superfamily: heat-stable enterotoxin ST  
C:Keywords: enterotoxin; heat-stable protein  
F:1-53/DNA: signal sequence and propeptide #status predicted <SIG>  
F:54-72/Product: heat-stable enterotoxin S-Ib #status experimental <MAT>  
F:59-64,60-68,63-71/Disulfide bonds: #status experimental

Query Match 12.9%; Score 70.5; DB 1; Length 72;  
Best Local Similarity 32.8%; Pred. No. 2.5; Matches 19; Conservative 6; Mismatches 18; Indels 15; Gaps 2;

Oy 63 YNPALPLDLPVCASQEAATFKALRTIATDE-----CELCINVACTGC 106  
::|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Db 15 FSP-FPQAKPVESSEKITLEKKNIATKNKSGPESMNSNYCCELCCNPACTGC 71  
::|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 10  
A82249 response regulator VC1050 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
C:Species: Vibrio cholerae  
C:Date: 18-Aug-2000 #aequation\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C:Accession: A82249  
R:Heidelberger, J.F.; Eissen, J.A.; Nelson, W.C.; Clayton, R.A.; Winn, M.L.; Dodson, R.J.;  
Charadson, D.; Ermolaeva, M.D.; Vamathevan, J.; Baas, S.; Qin, H.; Dragoi, I.; Sellers, P.  
L. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833; PMID:10952301  
A:Accession: A82249  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-368 <HBI>  
A:Cross-references: UNIPROT:Q9KT55; GB:AE004186; GB:AE003852; NID:g9655507; PIDN:AAF9420;  
A:Experimental source: serogroup O1, strain N16961; biotype El Tor  
C:Genetics:  
A:Map position: 1  
A:Gene: VC1050  
C:Superfamily: response regulator, Hnr type; response regulator homology

Query Match 12.9%; Score 70.5; DB 2; Length 368;  
Best Local Similarity 30.7%; Pred. No. 14; Matches 31; Conservative 21; Mismatches 30; Indels 19; Gaps 6;

Oy 4 SOLWAVALILLV-----LQAQGVYIKYHFQV-----QLSVKLNLEBK-----QMS 48  
::|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Db 19 AQITAAKLILLVDPPFRMTKG-YLKGYQYIEAENGLDKLRDRHPDLILCDLS 77  
::|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Oy 49 DPQQSKGLLPDVCLN-PALPLDLPVCASQEAATFKALR 88  
::|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Db 78 MPVLNGIEFVEEVSSVPSPPLVV-VSATEEMADVAKLR 116  
::|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 11  
SI8786  
N:Xin type A precursor (allele 2) - Streptococcus pyogenes phage (strain MGAS250 isolate  
N:Alternate names: scarlet fever toxin  
C:Species: Streptococcus pyogenes phage  
A:Variety: strain MGAS250 isolate California; strain MGAS251 isolate California; strain  
isolate United Kingdom; strain MGAS496 isolate Germany  
C:Date: 29-Jan-1993 #sequence revision 29-Jan-1993 #text\_change 16-Jul-1999  
C:Accession: SI8786; SI8787; SI8788; SI8790; SI8792; SI8795; SI8799

AGRT  
agrin - rat

C;Species: Rattus norvegicus (Norway rat)  
C;Date: 31-Mar-1999 #sequence\_revision 31-Mar-1993 #text\_change 09-Jul-2004  
C;Accession: JH0399; A38856  
R;Rupp, F.; Payan, D.G.; Magill-Solc, C.; Cowan, D.M.; Scheller, R.H.  
Neuron 6, 811-823, 1991

A>Title: Structure and expression of a rat agrin.  
A;Reference number: JH0399; MUID:91222570; PMID:1851019  
A;Accession: JH0399  
A;Molecule type: mRNA  
A;Residues: 1-1779;1799-1959 <RUP>  
A;Cross-references: UNIPROT:P25304; GB:M64780; NID:g202798; PIDN:AAA40703.1; PID:g202800  
A;Experimental source: embryonic spinal cord  
A>Note: it is uncertain whether Met-1, Met-18, or Met-24 is the initiator  
R;Rupp, F.; Oezcelik, T.; Linnal, M.; Peterson, K.; Francke, U.; Scheller, R.  
J. Neurosci. 12, 3535-3544, 1992

A>Title: Structure and chromosomal localization of the mammalian agrin gene.  
A;Reference number: A38856; MUID:92407628; PMID:1326608  
A;Accession: A38856  
A;Molecule type: mRNA  
A;Residues: 1780-1798 <RU2>  
A;Cross-references: GB:S44194  
C;Comment: This protein mediates the motor neuron-induced aggregation of acetylcholine x yCholine receptor clustering activity.

C;Superfamily: agrin; EGF homology; Kazal protease inhibitor homology; laminin G repeat  
F;1-1959/Product: agrin, form 1 #status predicted <AGI>  
F;1-1787,1799-1959/Product: agrin, form 4 #status predicted <AG4>  
F;1-1779,1799-1959/Product: agrin, form 3 #status predicted <AG3>  
F;1-1779,1788-1959/Product: agrin, form 5 #status predicted <AG5>  
F;1-1143,1153-1959/Product: agrin, form 2 #status predicted <AG2>  
F;22-50/Region: hydrophobic  
F;88-137/Domain: Kazal protease inhibitor homology <KPI1>  
F;163-212/Domain: Kazal protease inhibitor homology <KPI2>  
F;236-284/Domain: Kazal protease inhibitor homology <KPI3>  
F;307-356/Domain: Kazal protease inhibitor homology <KPI4>  
F;381-429/Domain: Kazal protease inhibitor homology <KPI5>  
F;446-494/Domain: Kazal protease inhibitor homology <KPI6>  
F;511-559/Domain: Kazal protease inhibitor homology <KPI7>  
F;540-542/Region: motor neuron attachment (L-R-E) motif  
F;596-645/Domain: Kazal protease inhibitor homology <KPI8>  
F;688-739/Domain: laminin-type EGF-like homology <LE1>  
F;742-786/Domain: laminin-type EGF-like homology <LE2>  
F;814-864/Domain: Kazal protease inhibitor homology <KPI9>  
F;869-992/Region: Kazal/threonine-rich  
F;1084-1085/Region: motor neuron attachment (L-R-E) motif  
F;1147-1215/Domain: serine/threonine-rich  
F;1224-1257/Domain: EGF homology <EG1>  
F;1287-1442/Domain: laminin G repeat homology <LG1>  
F;1444-1476/Domain: EGF homology <EG2>  
F;1483-1515/Domain: EGF homology <EG3>  
F;1555-1706/Domain: laminin G repeat homology <LG2>  
F;1713-1747/Domain: EGF homology <EG4>  
F;1807-1959/Domain: laminin G repeat homology <LG3>  
F;97-116,105-137,171-191,180-212,244-263,252-284,316-335,324-356,389-408,397-429,454-473  
-1476,1483-1494,1488-1504,1506-1515/Disulfide bonds: #status predicted  
F;145,672,827,957/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 12.7%; Score 69.5; DB 1; Length 1959;  
Best Local Similarity 26.2%; Pred. No. 1e+02;  
Matches 21; Conservative 7; Mismatches 19; Indels 33; Gaps 3;

Gy 49 DPQQSGLLPDVCYNPALPLDLQPVCSAQEAASTFKALRTATDECEL----- 97  
: : : : : : :  
Db 79 EPSVEDPGRASCVKCKGNCAPTAVPCGSD--ASTY-----SNEELQAQCNOQRRI 129  
: : : : : : :  
Gy 98 -----CINVACT 104  
: : : : : : :

Dd 130 RLRLQGPGSRDPFCANVTGS 149

Search completed: March 26, 2005, 16:49:40  
Job time : 23.438 secs

RESULT 15  
\$18783



GenCore version 5.1.6

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2005, 16:37:18 ; Search time 91.2993 Seconds  
(without alignments)  
594.533 Million cell updates/sec

Title: US-10-775-481A-55

Perfect score: 548

Sequence: 1 MSGSQLWAVLLLLVLSAQ.....LRTIATDELCINACTGC 106

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 548   | 100.0       | 106    | 1     | GUAAU RAT   |
| 2          | 500.5 | 91.3        | 107    | 2     | Q8R5G8      |
| 3          | 500   | 91.2        | 106    | 1     | GUAAU MOUSE |
| 4          | 500   | 91.2        | 106    | 2     | Q9QU03      |
| 5          | 356   | 65.0        | 112    | 1     | GUAAU HUMAN |
| 6          | 335   | 61.1        | 111    | 1     | GUAAU CAVPO |
| 7          | 311   | 56.8        | 113    | 1     | GUAAU PIG   |
| 8          | 278.5 | 50.8        | 109    | 1     | GUAAU DIDMA |
| 9          | 174.5 | 31.8        | 109    | 1     | GUAAU PIG   |
| 10         | 167.5 | 30.6        | 108    | 2     | Q98T10      |
| 11         | 167.5 | 30.6        | 108    | 2     | Q7Z2S0      |
| 12         | 167.5 | 30.6        | 116    | 2     | Q98T19      |
| 13         | 166.5 | 30.4        | 115    | 1     | GUAAU HUMAN |
| 14         | 150.5 | 27.5        | 110    | 2     | Q7Z2S1      |
| 15         | 149   | 27.2        | 107    | 1     | GUAAU CAVPO |
| 16         | 142   | 25.9        | 109    | 2     | Q7Z2S2      |
| 17         | 137   | 25.0        | 115    | 2     | Q8R5G9      |
| 18         | 135   | 24.6        | 115    | 1     | GUAAU RAT   |
| 19         | 134.5 | 24.5        | 116    | 1     | GUAAU MOUSE |
| 20         | 99    | 18.1        | 119    | 2     | Q70674      |
| 21         | 75    | 13.7        | 61     | 2     | Q5VEG8      |
| 22         | 74    | 13.5        | 354    | 1     | LXAL PHOLE  |
| 23         | 73.5  | 13.4        | 346    | 2     | Q943F0      |
| 24         | 73.5  | 13.4        | 397    | 2     | Q8S255      |
| 25         | 73    | 13.3        | 189    | 1     | COPI2 YEAST |
| 26         | 73    | 13.3        | 254    | 2     | Q5VFN9      |
| 27         | 73    | 13.3        | 1139   | 1     | M2A2 HUMAN  |
| 28         | 72.5  | 13.2        | 228    | 2     | Q26057      |
| 29         | 72    | 13.1        | 800    | 2     | Q6DRN3      |
| 30         | 71    | 13.0        | 268    | 2     | Q7S0G1      |
| 31         | 71    | 13.0        | 713    | 2     | Q6L968      |

32 71 13.0 775 2 Q72HT2 thermus the  
33 70.5 12.9 368 2 Q9KT55 Q9kt55 vibrio chol  
34 70.5 12.9 710 2 Q94U06 Q94u06 arabidopsis  
35 70.5 12.9 710 2 Q8H1E8 Q8h1e8 arabidopsis  
36 70.5 12.9 751 2 Q9LUU0 Q9luu0 arabidopsis  
37 70 12.8 72 1 HST2\_ECOLI Q47185 escherichia  
38 70 12.8 274 2 Q8ZL88 Q8z188 salmonella  
39 70 12.8 793 2 Q8IBJ8 Q8ibj8 plasmodium  
40 69.5 12.7 236 2 Q57453 Q57453 streptococ  
41 69.5 12.7 300 2 Q9RC89 Q9rc89 bacillus ha  
42 69.5 12.7 458 2 Q9KB34 Q9kb34 bacillus ha  
43 69.5 12.7 610 2 Q8CM38 Q8cm38 streptococ  
44 69.5 12.7 639 2 Q8ESL4 Q8es14 streptococ  
45 69.5 12.7 824 2 Q90778 Q90778 gallus galli

## ALIGNMENTS

## RESULT 1

GUAAU RAT STANDARD; PRT; 106 AA.  
AC P70658;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Uroguanylin precursor (UGN) (Guanylate cyclase activator 2B).  
GN Name=Guca2b;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 92-106.  
RC STRAIN=Sprague-Dawley;  
RX MEDLINE=97248740; PubMed=9094754; DOI=10.1016/S0167-0115(96)02103-9;  
RA Li Z., Perkins A.G., Peters M.F., Campa M.J., Goy M.F.;  
RT "Purification, cDNA sequence, and tissue distribution of rat uroguanylin."  
RL Regul. Pept. 68:45-56(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97131589; PubMed=8977100; DOI=10.1016/S0014-5793(96)01235-5;  
RA Miyazato M., Nakazato M., Matsukura S., Kangawa K., Matsuo H.;  
RT "Uroguanylin gene expression in the alimentary tract and extra-gastrointestinal tissues".  
RL FEBS Lett. 398:170-174(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley; TISSUE=Small intestine;  
RX MEDLINE=97319300; PubMed=9176203;  
RA Blanchard R.K., Cousins R.J.;  
RT "Upregulation of rat intestinal uroguanylin mRNA by dietary zinc restriction".  
RL Am. J. Physiol. 272:G972-G978(1997).  
CC -1- FUNCTION: Endogenous activator of intestinal guanylate cyclase. It stimulates this enzyme through the same receptor binding region as the heat-stable enterotoxins. May be a potent physiological regulator of intestinal fluid and electrolyte transport. May be an autocrine/paracrine regulator of intestinal salt and water transport.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Expressed not only in the gastrointestinal tract but also in the lung, pancreas and kidney.  
CC -1- SIMILARITY: Belongs to the guanylin family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

```
-----
EMBL; U73898; AAB18331.1; -
EMBL; U41322; AAB18760.1; -
EMBL; U75186; AAB61209.1; -
HSSP; Q16661; IUYA.
RSD; 620044; Guca2b.
InterPro; IPR000879; Guanylin.
PIRSF; PIRSF001849; Guanylin; 1.
PRINTS; PR00774; GUANYLIN.
ProDom; PD005588; Guanylin; 1.
Direct protein sequencing; Signal.
FT SIGNAL 1 21 Potential.
FT PROPEP 22 91
FT PEPTIDE 92 106 Uroguanylin.
FT DISULFID 62 75 Potential.
FT DISULFID 95 103 By similarity.
FT DISULFID 98 106 By similarity.
SQ SEQUENCE 106 AA; 11573 MW; 9FB5F88A9B1DD077 CRC64;

Query Match 100.0%; Score 548; DB 1; Length 106;
Best Local Similarity 100.0%; Pred. No. 5e-48;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGSQLAAVLLLVLSAQGVYIKYHGFQVQLSVKLNLEEKQWSDPQQKSGLLPD 60
D 1 MSGSQLAAVLLLVLSAQGVYIKYHGFQVQLSVKLNLEEKQWSDPQQKSGLLPD 60
QY 61 VCNYPALPLDLPQVCASQEAATFKALRTIATDECELCINVACTGC 106
D 61 VCNYPALPLDLPQVCASQEAATFKALRTIATDECELCINVACTGC 106

RESULT 2
Q8R5G8 PRELIMINARY; PRT; 107 AA.
AC Q8R5G8;
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 21, Last annotation update)
DE Uroguanylin.
OS Notomys alexis (Spinifex hopping mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Notomys.
OX NCBI_TaxID=184396;
RN [1]
SEQUENCE FROM N.A.
RA Donald J.A., Bartolo R.C.;
RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF469496; AAL77417.1; -
DR HSSP; Q16661; IUYA.
DR GO; GO:0008047; F1enzyme activator activity; IEA.
DR InterPro; IPR000879; Guanylin.
DR Pfam; PF02058; Guanylin; 1.
DR PIRS; PIRSF001849; Guanylin; 1.
DR PRINTS; PR00774; GUANYLIN.
DR ProDom; PD005588; Guanylin; 1.
SQ SEQUENCE 107 AA; 11618 MW; 735110CACB60DA97 CRC64;

Query Match 91.3%; Score 500.5; DB 2; Length 107;
Best Local Similarity 90.7%; Pred. No. 3.5e-43;
Matches 97; Conservative 7; Mismatches 2; Indels 1; Gaps 1;

QY 1 MSGSQLAAVLLLVLSAQGVYIKYHGFQVQLSVKLNLEEKQWSDPQQKSG-LLP 59
D 1 MSGSQLAAVLLLVLSAQGVYIKYHGFQVQLSVKLNLEEKQWSDPQQKSGLLP 60
QY 60 DVCPNLPALPLDLPQVCASQEAATFKALRTIATDECELCINVACTGC 106
D 61 DVCPNLPALPLDLPQVCASQEAATFKALRTIATDECELCINVACTGC 107

RESULT 3
```

```
GUAN_MOUSE STANDARD; PRT; 106 AA.
ID GUAN_MOUSE
AC Q09051;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Uroguanylin precursor (UGN) (Guanylate cyclase activator 2B).
GN Name-Guca2b, (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=97434109; PubMed=9287995;
RX Whitaker T.L., Witte D.P., Scott M.C., Cohen M.B.;
RA "Uroguanylin and guanylin: distinct but overlapping patterns of
RT messenger RNA expression in mouse intestine.";
RL Gastroenterology 113:1000-1006(1997).
RN [2]
REVISION TO 17.
RA Sanford L.P., Cohen M.B.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Endogenous activator of intestinal guanylate cyclase. It
CC stimulates this enzyme through the same receptor binding region as
CC the heat-stable enterotoxins. May be a potent physiological
CC regulator of intestinal fluid and electrolyte transport. May be an
CC autocrine/paracrine regulator of intestinal salt and water
CC transport (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Localized predominantly in intestinal villi
CC and the corticomedullary junction of the kidney.
CC -!- SIMILARITY: Belongs to the guanylin family.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; U95182; AAB82750.2; -
DR EMBL; U90727; AAB53314.1; -
DR HSSP; Q16661; IUYA.
DR MGD; MGI:1270851; Guca2b.
DR InterPro; IPR000879; Guanylin.
DR Pfam; PF02058; Guanylin; 1.
DR PRINTS; PR00774; GUANYLIN.
DR ProDom; PD005588; Guanylin; 1.
KW Signal.
FT SIGNAL 1 21 Potential.
FT PROPEP 22 91 Uroguanylin.
FT PEPTIDE 92 106 Potential.
FT DISULFID 62 75 By similarity.
FT DISULFID 95 103 By similarity.
FT DISULFID 98 106 By similarity.
FT CONFLICT 17 17 Q -> R (in Ref. 1; AAB53314).
SQ SEQUENCE 106 AA; 11627 MW; 30FF1CCE9D293DA8 CRC64;

Query Match 91.2%; Score 500; DB 1; Length 106;
Best Local Similarity 90.6%; Pred. No. 3.9e-43;
Matches 96; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSGSQLAAVLLLVLSAQGVYIKYHGFQVQLSVKLNLEEKQWSDPQQKSGLLPD 60
D 1 MSGSQLAAVLLLVLSAQGVYIKYHGFQVQLSVKLNLEEKQWSDPQQKSGLLPD 60
QY 61 VCNYPALPLDLPQVCASQEAATFKALRTIATDECELCINVACTGC 106
D 61 VCNYPALPLDLPQVCASQEAATFKALRTIATDECELCINVACTGC 106
```

|                       |   |
|-----------------------|---|
| RL                    | Nature 420:563-573 (2002).  |
| RN                    | [8]   |
| RP                    | SEQUENCE FROM N.A.  |
| RC                    | STRAIN=C57BL/6J; TISSUE=Kidney;   |
| RD                    | MEDLINE=20493374; PubMed=11042159; DOI=10.1101/gr.145100;   |
| RX                    | Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  |
| RA                    | Kanno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;   |
| RA                    | "Normalization and subtraction of cap-trapper-selected cDNAs to   |
| RT                    | prepare full-length cDNA libraries for rapid discovery of new genes.";  |
| RL                    | Genome Res. 10:1617-1630(2000).   |
| RN                    | [9]   |
| RP                    | SEQUENCE FROM N.A.  |
| RC                    | STRAIN=C57BL/6J; TISSUE=Kidney;   |
| RD                    | MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;   |
| RX                    | Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,   |
| RA                    | Kanno H., Ishiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,  |
| RA                    | Sumi N., Akai Y., Nakamura S., Hazama M., Nishino T., Harada A.,  |
| RA                    | Yamanoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  |
| RA                    | Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,  |
| RA                    | Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuda S., Kawai J.,  |
| RA                    | Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  |
| RT                    | "RIKEN integrated sequence analysis (RISA) system-394-format  |
| RT                    | sequencing pipeline with 384 multicapillary sequencer.";  |
| RL                    | Genome Res. 10:1757-1771(2000).   |
| RN                    | [10]  |
| RP                    | SEQUENCE FROM N.A.  |
| RC                    | STRAIN=C57BL/6J; TISSUE=Kidney;   |
| RD                    | Adachi J., Aizawa K., Akahira S., Akimura T., Atai A., Aono H.,   |
| RA                    | Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,   |
| RA                    | Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,   |
| RA                    | Inotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato C.,  |
| RA                    | Kawai J., Koijima Y., Konno H., Kouda M., Koya S., Kurihara H.,   |
| RA                    | Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,   |
| RA                    | Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,   |
| RA                    | Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,   |
| RA                    | Sogabe Y., Suzuki H., Tagami M., Tagawa Y., Takahashi P., Tanaka T.,  |
| RA                    | Tajima Y., Toyota T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  |
| RA                    | Muramatsu M., Hayashizaki Y.;   |
| RT                    | Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.   |
| RL                    | EMBL; AF006668; AAD09311.1; -   |
| DR                    | EMBL; BC024373; AJH24373.1; -   |
| DR                    | EMBL; U67800; AAD09215.1; -   |
| DR                    | EMBL; AK002364; BAB22042.1; -   |
| DR                    | HSSP; Q16661; 1UYA.   |
| DR                    | MGD; MG1:1270851; Guca2b.   |
| DR                    | GO; GO:0005615; Cytoplasmic space; TAS.   |
| DR                    | GO; GO:0006182; Protein biosynthesis; IMP.  |
| DR                    | GO; GO:0007589; Protein secretion; IMP.   |
| DR                    | GO; GO:0045776; Negative regulation of blood pressure; IMP.   |
| DR                    | InterPro; IPRO00879; Guanlylin.   |
| DR                    | Pfam; PF02058; Guanlylin; 1.  |
| DR                    | PIRSF; PIRSF001849; Guanlylin; 1.   |
| DR                    | PRINTS; PR00774; GUANYLIN.  |
| DR                    | PRODOM; PD005588; Guanlylin; 1.   |
| DR                    | SEQUENCE 106 AA; 11627 MW; 30FFICE9D293DA8 CRC64;   |
| SQ                    |   |
| Query Match           | 91.2%; Score 500; DB 2; Length 106;   |
| Best Local Similarity | 90.6%; Pred. No. 3 9e+43;   |
| Matches               | 96; Conservative 7; Mismatches 3; Indels 0; Gaps 0;   |
| QY                    | 1 MGSQSLAAVLLLVLSAQGVYKYHGQVLSEVKKLNELEEKQMSPPQQSGLLPD 60<br>          :           :           :           :           :  |
| Db                    | 1 MSRSQSLAAVLLLVLSAQGVYKYHGQVLSEVKKLNELEEKQMSPPQQSGLLPA 60<br>          :           :           :           :           : |
| QY                    | 61 VCNPALPLDLQPVCSQAASSTFKALRTIATDECELCINACTGC 106<br>          :           :           :           :           :         |
| Db                    | 61 VCHNPALPLDLQPVCSQAASSTFKALRTIATDECELCINACTGC 106<br>          :           :           :           :           :        |
| RESULT 5              |   |
| GUAR_HUMAN            | STANDARD; PRT; 112 AA.  |
| ID                    | GUAR_HUMAN  |
| AC                    | Q16661;   |



CC autocrine/paracrine regulator of intestinal salt and water  
 CC transport.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the guanylin family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC  
 DR EMBL; Z74738; CAA98994.1; -.  
 DR HSPF; Q16661; IUYA.  
 DR InterPro; IPR000879; Guanylin.  
 DR Pfam; PF02058; Guanylin; 1.  
 DR PIRSF; PIRSF001849; Guanylin; 1.  
 DR PRINTS; PR00774; GUANYLIN.  
 DR ProDom; PD005588; Guanylin; 1.  
 KW SIGNAL.  
 FT SIGNAL 1 26 Potential.  
 FT PROPEP 27 96  
 FT PEPTIDE 97 111 Uroguanylin.  
 FT DISULFID 67 80 Potential.  
 FT DISULFID 100 108 By similarity.  
 FT DISULFID 103 111 By similarity.  
 SQ SEQUENCE 111 AA; 12125 MW; 7C3366A721FE0411 CRC64;  
 Query Match 61.1%; Score 335; DB 1; Length 111;  
 Best Local Similarity 65.3%; Pred. No. 2.8e-26;  
 Matches 64; Conservative 13; Mismatches 21; Indels 0; Gaps 0;  
 QY 9 AVLLLLVLSAQGVYIKYGHFQVLESVKKLEELKEKMSDPQOKSGLLPDVCYNPALP 68  
 Db 14 AVLLLLVLSAQGVYIKYGHFQVLESVKKLEELKEKMSDPQOKSGLLPDVCYNPALP 73  
 QY 69 LDLPQVCASQAASSTFKALRTIATDECELCINACTGC 106  
 Db 74 LDLPQVCASQAASSTFKALRTIATDECELCINACTGC 111  
 RESULT 7  
 GUAY\_PIG STANDARD; PRT; 113 AA.  
 AC O13009;  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Uroguanylin precursor (UGN) (Guanylate cyclase activator 2B)  
 GN Name=GUCA2B;  
 OS Sus scrofa (pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Jejunum;  
 RA Hill O., Maegert H.J., Forssmann W.-G.;  
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Endogenous activator of intestinal guanylate cyclase. It  
 CC stimulates this enzyme through the same receptor binding region as  
 CC the heat-stable enterotoxins. May be a potent physiological  
 CC regulator of intestinal fluid and electrolyte transport. May be an  
 CC autocrine/paracrine regulator of intestinal salt and water  
 CC transport (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).  
 CC -!- SIMILARITY: Belongs to the guanylin family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC  
 DR EMBL; Z83746; CAB06042.1; -.  
 DR HSPF; Q16661; IUYA.  
 DR InterPro; IPR000879; Guanylin.  
 DR Pfam; PF02058; Guanylin; 1.  
 DR PIRSF; PIRSF001849; Guanylin; 1.  
 DR PRINTS; PR00774; GUANYLIN.  
 DR ProDom; PD005588; Guanylin; 1.  
 KW SIGNAL.  
 FT SIGNAL 1 27 Potential.  
 FT PROPEP 28 89  
 FT PEPTIDE 90 113 GCAP-II.  
 FT PEPTIDE 98 113 Uroguanylin.  
 FT DISULFID 68 81 By similarity.  
 FT DISULFID 101 109 By similarity.  
 FT DISULFID 104 112 By similarity.  
 SQ SEQUENCE 113 AA; 12044 MW; 8160573287BAB642 CRC64;  
 Query Match 56.8%; Score 311; DB 1; Length 113;  
 Best Local Similarity 59.2%; Pred. No. 7.9e-22;  
 Matches 58; Conservative 18; Mismatches 24; Indels 0; Gaps 0;  
 QY 9 AVLLLLVLSAQGVYIKYGHFQVLESVKKLEELKEKMSDPQOKSGLLPDVCYNPALP 68  
 Db 15 ALVFLVLLQGTQSVYIQYQGFVQLKSVKLLSDLEGWAPSPRLQAQSPQSVCHSALP 74  
 QY 69 LDLPQVCASQAASSTFKALRTIATDECELCINACTGC 106  
 Db 75 PDLQPICTQSEAASITQALRTIATDECELCINACTGC 112  
 RESULT 8  
 GUAY\_DIDMA STANDARD; PRT; 109 AA.  
 AC Q28358;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Uroguanylin precursor (UGN) (Guanylate cyclase activator 2B).  
 GN Name=GUCA2B;  
 OS Didelphis marsupialis virginiana (North American opossum).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.  
 OX NCBI\_TaxID=9267;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Colon;  
 RA Fan X., Hamra F.K., Freeman R.H., Eber S.L., Krause W.J., Lim R.W.,  
 RA Pace V.M., Currie M.G., Forte L.R.;  
 RT "Uroguanylin: cloning of preproguanylin cDNA, mRNA expression in  
 RT the intestine and heart and isolation of uroguanylin and  
 RT prouroguanylin from plasma";  
 RL Biochem. Biophys. Res. Commun. 219:457-462 (1996).  
 RN [2]  
 RP SEQUENCE OF 95-109.  
 RC TISSUE=Urine;  
 RA Hamra F.K., Forte L.R., Eber S.L., Pithorodeckyj N.V., Krause W.J.,  
 RA Freeman R.H., Chin D.T., Tompkins J.A., Fox K.F., Smith C.E.,  
 RA Duffin K.L., Siegel N.R., Currie M.G.;  
 RT "Uroguanylin: structure and activity of a second endogenous peptide  
 RT that stimulates intestinal guanylate cyclase";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:10464-10468 (1993).  
 CC -!- FUNCTION: Endogenous activator of intestinal guanylate cyclase. It  
 CC stimulates this enzyme through the same receptor binding region as  
 CC the heat-stable enterotoxins. May be a potent physiological  
 CC regulator of intestinal fluid and electrolyte transport. May be an  
 CC autocrine/paracrine regulator of intestinal salt and water

CC transport.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Small and large intestine and atria and  
CC ventricles of heart. Both uroguanylin and prouroguanylin are found  
CC in plasma.  
CC -!- SIMILARITY: Belongs to the guanylin family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: U49353; AAB00553.1; -.  
CC HSSP: Q16661; IUYA.  
DR InterPro: IPR000879; Guanylin.  
DR Pfam: PF02058; Guanylin; 1.  
DR PIRSF: PIRSF001849; Guanylin; 1.  
DR PRINTS: PR00774; GUANYLIN.  
DR ProDom: PD005588; Guanylin; 1.  
KW Direct protein sequencing; Signal.  
FT SIGNAL 1 23 Potential.  
FT PROPEP 24 94 Uroguanylin.  
FT PEPTIDE 95 109 Potential.  
FT DISULFID 65 78 By similarity.  
FT DISULFID 98 106 By similarity.  
FT DISULFID 101 109 By similarity.  
SQ SEQUENCE 109 AA; 12040 MW; AE948E210CA3AE7A CRC64;  
  
Query Match 50.8%; Score 278.5; DB 1; Length 109;  
Best Local Similarity 55.8%; Pred. No. 1.6e-20;  
Matches 55; Conservative 20; Mismatches 23; Indels 1; Gaps 1;  
  
QY 9 AVLLLLVLSQAQGVYIKYHGFQVQLSVKLNLESEKQMSDPQQQKSGLLPDV-CYNPAL 67  
DB 11 AMLLVLAQNTQSVYIQYEGFQVQLSVKLNLESEKQMSDPQQQKSGLLPDV-CYNPAL 70  
  
QY 68 PDLQPVCAQSAASTFKALRTIATDECELCINVACTGC 106  
DB 71 PSLDQPVCAQSAANIFRALRSISQBDCELCINVACTGC 109  
  
RESULT 9  
GUAN\_PIG STANDARD; PRT; 109 AA.  
AC GUAN\_PIG  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Guanylin precursor (Guanylate cyclase activator 2A).  
GN Name=GUCA2A; Synonyms=GUCA2;  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99268812; PubMed=10334930; DOI=10.1006/bbrc.1999.0719;  
RA Maegert H.J., Hill O., Zucht H.D., Martin S., Meyer M.,  
RA Forssmann W.-G., Adermann K.;  
RT "Porcine guanylin and uroguanylin: cDNA sequences, deduced amino acid  
RT sequences, and biological activity of the chemically synthesized  
RT peptides."  
RL Biochem. Biophys. Res. Commun. 259:141-148(1999).  
CC -!- FUNCTION: Endogenous activator of intestinal guanylate cyclase. It  
CC stimulates this enzyme through the same receptor binding region as  
CC the heat-stable enterotoxins.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: Belongs to the guanylin family.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: Z73607; CAA97974.1; -.  
CC HSSP: Q02747; IGNA.  
DR InterPro: IPR000879; Guanylin.  
DR Pfam: PF02058; Guanylin; 1.  
DR PIRSF: PIRSF001849; Guanylin; 1.  
DR PRINTS: PR00774; GUANYLIN.  
DR ProDom: PD005588; Guanylin; 1.  
KW Signal.  
FT SIGNAL 1 21 Potential.  
FT PROPEP 22 94  
FT PEPTIDE 95 109 Guanylin.  
FT DISULFID 98 106 By similarity.  
FT DISULFID 101 109 By similarity.  
SQ SEQUENCE 109 AA; 11938 MW; 6368BD1FB0B07A74 CRC64;  
  
Query Match 31.8%; Score 174.5; DB 1; Length 109;  
Best Local Similarity 37.5%; Pred. No. 6.3e-10;  
Matches 39; Conservative 16; Mismatches 38; Indels 11; Gaps 3;  
  
QY 6 LWAVALLLVLSQAQGVYIKYHGFQVQLSVKLNLESEKQMSDPQQQKSGLLPDV-CYN 64  
DB 14 VMAAL-----AGGVTKDGEFSLSVSKKLQDLQELQKPRNRLDGPPIPVLCNS 65  
  
QY 65 PALPDLQPVCAQSAASTFKALRTIATDECELCINVACTGC 106  
DB 66 PKFPPELPICQKNABEILERLETIAQDPSTCEICAYAAACGC 109  
  
RESULT 10  
Q98TIO PRELIMINARY; PRT; 108 AA.  
AC Q98TIO  
DT 01-JUN-2001 (TREMELrel. 17, Created)  
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)  
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)  
DE Guanylin.  
GN Name=GUCA2I;  
OS Anguilla anguilla (European freshwater eel).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;  
OC Anguilla.  
OX NCBI\_TaxID=7936;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21139737; PubMed=11243845; DOI=10.1006/bbrc.2001.4485;  
RA Comrie M.M., Cutler C.P., Cramb G.;  
RT "Cloning and Expression of Guanylin from the European eel (Anguilla  
RT anguilla)."  
RL Biochem. Biophys. Res. Commun. 281:1078-1085(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX Comrie M.M.;  
RL Thesis (2000), Department of School of Biology, University of St  
RL Andrews, St Andrews, United Kingdom.  
DR EMBL: AJ301672; CAC35448.1; -.  
DR HSSP: Q02747; 108R.  
DR GO: GO:0008047; F:enzyme activator activity; IEA.  
DR InterPro: IPR006058; 2Fe2S fd BS.  
DR InterPro: IPR000879; Guanylin.  
DR Pfam: PF02058; Guanylin; 1.  
DR PIRSF: PIRSF001849; Guanylin; 1.  
DR PRINTS: PR00774; GUANYLIN.  
DR ProDom: PD005588; Guanylin; 1.  
DR PROSITE: PS00197; 2FE2S\_FERREDOXIN; UNKNOWN 1.  
SQ SEQUENCE 108 AA; 11584 MW; 8A3B6D490E7C858D CRC64;

[illegible]

RA MEDLINE=95199289; PubMed=7892222;  
RA Hill O., Kuhn M., Zucht H.D., Cetin Y., Kulaksiz H., Adermann K.,  
RA Klock G., Reckemmer G., Forssmann W.-G., Magert H.J.;  
RT "Analysis of the human guanylin gene and the processing and cellular  
RT localization of the peptide";  
RL Proc. Natl. Acad. Sci. U.S.A. 92:2046-2050(1995).  
RN [4]  
RN SEQUENCE OF 22-68.  
RP MEDLINE=93178628; PubMed=8095028; DOI=10.1016/0014-5793(93)80022-M;  
RX Kuhn M., Raida M., Adermann K., Schulz-Knappe P., Gerzer R.,  
RA Heim J.-W., Forssmann W.-G.;  
RT "The circulating bioactive form of human guanylin is a high molecular  
RT weight peptide (10.3 kDa).";  
RL FEBS Lett. 318:205-209(1993).  
RN [5]  
RP STRUCTURE BY NMR OF 101-115.  
RX MEDLINE=95034794; PubMed=7947768;  
RA Skelton N.J., Garcia K.C., Goeddel D.V., Quan C., Burnier J.P.;  
RT "determination of the solution structure of the peptide hormone  
RT guanylin: observation of a novel form of topological  
RT stereoisomerism";  
RL Biochemistry 33:13581-13592(1994).  
CC -!- FUNCTION: Endogenous activator of intestinal guanylate cyclase. It  
CC stimulates this enzyme through the same receptor binding region as  
CC the heat-stable enterotoxins.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Highly expressed in ileum and colon. Found in  
CC plasma.  
CC -!- SIMILARITY: Belongs to the guanylin family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M97496; AAA35915.1; -;  
DR EMBL; M95174; AAAS8625.1; -;  
DR EMBL; X74322; CAC22258.1; -;  
DR PIR; A46279; A46279.  
DR PDB; 1GNA; NMR; @=103-115.  
DR PDB; 1GNB; NMR; @=103-115.  
DR PDB; 1ORR; NMR; A=22-115.  
DR Genew; HGNC:4682; GUC2A.  
DR MIM; 133392; -;  
DR GO; GO:0005576; C:extracellular; NAS.  
DR GO; GO:0030250; F:guanylate cyclase activator activity; NAS.  
DR GO; GO:0003179; F:hormone activity; NAS.  
DR GO; GO:0007165; P:signal transduction; NAS.  
DR InterPro; IPR000879; Guanylin.  
DR Pfam; PF02058; Guanylin; 1.  
DR PIRSF; PIRSF001849; Guanylin; 1.  
DR PRINTS; PR00774; GUANYLIN.  
DR ProDom; PD005588; Guanylin; 1.  
KW 3D-structure; Direct protein sequencing; Signal.  
FT SIGNAL 1 21  
FT CHAIN 22 115 HMW-guanylin.  
FT PEPTIDE 101 115 Guanylin.  
FT DISULFID 104 112  
FT DISULFID 107 115  
FT TURN 104 107  
FT TURN 110 111  
SQ SEQUENCE 115 AA; 12448 MW; C644C03BAFFC26FA CRC64;  
Query Match 30.4%; Score 166.5; DB 1; Length 115;  
Best Local Similarity 35.1%; Pred. No. 4.4e-09;  
Matches 40; Conservative 15; Mismatches 32; Indels 27; Gaps 4;  
QY 7 WAAVLLLVLSQAGVYIKYHGFGVQLSVKLNLEBKQMSDPQOQKSG-----56  
DB 15 WAAV-----AGGVTVQDGNFSLSVKLNLEBKQMSDPQOQKSG-----POBFRVYKLNFAPIPG 61

QY 57 --LLPDVVCNFPALPDLQPVCAQEAASFTFKALRTIATD--ECBELCINVACTGC 106  
DB 62 EPVVILCSNPNFPPELKLPLCKEPNAQELQRLBEIAEDPGTCEICAVAACTGC 115  
RESULT 14  
Q7ZS1 PRELIMINARY; PRT; 110 AA.  
AC Q7ZS1;  
DT 01-JUN-2003 (TRENBLrel. 24, Created)  
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
DE Preprorenoguanilin.  
GN Name=renoguanilin;  
OS Anguilla japonica (Japanese eel).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;  
OC Anguilla.  
OX NCBI\_TaxID=7937;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Intestine;  
RX MEDLINE=22692502; PubMed=12684514; DOI=10.1074/jbc.M303111200;  
RA Yuge S., Inoue K., Hyodo S., Takei Y.;  
RT "A novel guanylin family (guanylin, uroguanylin, and renoguanilin) in  
RT eels: possible osmoregulatory hormones in intestine and kidney.";  
RL J. Biol. Chem. 278:22726-22733(2003).  
DR EMBL; AB080641; BAC76010.1; -;  
DR HSSP; Q02747; 108R.  
DR GO; GO:0008047; F:enzyme activator activity; IEA.  
DR InterPro; IPR006058; 2Fe2S fd BS.  
DR InterPro; IPR000879; Guanylin.  
DR Pfam; PF02058; Guanylin; 1.  
DR PIRSF; PIRSF001849; Guanylin; 1.  
DR PRINTS; PR00774; GUANYLIN.  
DR ProDom; PD005588; Guanylin; 1.  
DR PROSITE; PS00197; 2FE2S FERREDOXIN; UNKNOWN 1.  
SQ SEQUENCE 110 AA; 11646 MW; B161A03A913894F9 CRC64;  
Query Match 27.5%; Score 150.5; DB 2; Length 110;  
Best Local Similarity 39.2%; Pred. No. 1.8e-07;  
Matches 40; Conservative 15; Mismatches 38; Indels 9; Gaps 4;  
QY 10 VLLLVLSQAGVYIKYHGFGVQLSVKLNLEBKQMS---DPOQKSGLLPDVCYNPA 66  
DB 12 VALALVTEAVQ---VEEGSFPLDAVKKLELMGDMTVKQSPRLAKTSTTA-VCTNPD 67  
QY 67 LPDLQPVCAQEAASFTFKALRTIA--TDEBELCINVACTGC 106  
DB 68 LPAPVFLPLCKSKGAANSFFRLGFAARADLCICAFAACTGC 109  
RESULT 15  
GUAN\_CAVPO  
ID GUAN\_CAVPO STANDARD; PRT; 107 AA.  
AC P70664;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Guanylin precursor (Guanylate cyclase activator 2A).  
GN Name=GUC2A2; Synonyms=GUC2;  
OS Cavia porcellus (Guinea pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Caviidae; Cavia.  
OX NCBI\_TaxID=10141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Colon;  
RA Kruhoeffner M., Cetin Y., Kaempf U., Forssmann W.-G.;  
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Endogenous activator of intestinal guanylate cyclase. It  
CC stimulates this enzyme through the same receptor binding region as



```

CC the heat-stable enterotoxins.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the guanylin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z74736; CAA98991.1; -.
DR EMBL; Z74737; CAA98992.1; -.
DR HSSP; Q02747; IGNA.
DR InterPro; IPR000879; Guanylin.
DR Pfam; PF02058; Guanylin; 1.
DR PIRSF; PIRSF001849; Guanylin; 1.
DR PRINTS; PR00774; GUANYLIN.
DR PRODom; PD005588; Guanylin; 1.
KW Signal.
FT SIGNAL 1 20 Potential.
FT PROPEP 21 92
FT PEPTIDE 93 107 Guanylin.
FT DISULFID 96 104 By similarity.
FT DISULFID 99 107 By similarity.
SQ SEQUENCE 107 AA; 11483 MW; C4BE4CE609409F45 CRC64;

Query Match 27.2%; Score 149; DB 1; Length 107;
Best Local Similarity 33.3%; Pred. No. 2.5e-07;
Matches 34; Conservative 14; Mismatches 44; Indels 10; Gaps 2;

Qy 7 WAAVLLLLVLSAQGVYIKYHGQVQLSESVKKLNELEEKQMSDPQQQKGLLPDVCYNPA 66
Db 14 WAALVGAVTVQDGD-----FSPSLESVKKLQLEAPESKVGQRKFKVAPPLCSFSG 65

Qy 67 LPDLQPVCSAQEAATFKALRTIATD--ECELINVACTGC 106
Db 66 FPEELRPVCKEPNSQDILNRLAVIAQDPSTCEICAYAACAGC 107

```

Search completed: March 26, 2005, 16:48:35  
Job time : 92.2993 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2005, 16:35:58 ; Search time 122.628 Seconds  
(without alignments)  
353.241 Million cell updates/sec

Title: US-10-775-481A-56  
Perfect score: 579  
Sequence: 1 MGCRAAGLLPGVAVVLLLL.....RTIANDDELVCNVACTGCL 112

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 579   | 100.0       | 112    | 2     | AAW18498 Human GCA |
| 2          | 579   | 100.0       | 112    | 2     | AAW10595 Guanylate |
| 3          | 579   | 100.0       | 112    | 6     | ABG74823 Human uro |
| 4          | 579   | 100.0       | 112    | 8     | ADR45875 Amino aci |
| 5          | 554   | 95.7        | 108    | 2     | AAW18467 Human GCA |
| 6          | 554   | 95.7        | 108    | 2     | AAW18478 Human GCA |
| 7          | 554   | 95.7        | 108    | 2     | AAW23221 GCAP-II C |
| 8          | 445   | 76.9        | 84     | 2     | AAW18484 Human GCA |
| 9          | 445   | 76.9        | 84     | 2     | AAW23232 GCAP-II C |
| 10         | 429   | 74.1        | 81     | 2     | AAW18485 Human GCA |
| 11         | 429   | 74.1        | 81     | 2     | AAW23233 GCAP-II C |
| 12         | 412   | 71.2        | 78     | 2     | AAW18486 Human GCA |
| 13         | 412   | 71.2        | 78     | 2     | AAW23234 GCAP-II C |
| 14         | 407   | 70.3        | 77     | 2     | AAW18468 Human GCA |
| 15         | 407   | 70.3        | 77     | 2     | AAW18479 Human GCA |
| 16         | 407   | 70.3        | 77     | 2     | AAW23222 GCAP-II C |
| 17         | 389   | 67.2        | 73     | 2     | AAW18490 Human GCA |
| 18         | 389   | 67.2        | 73     | 2     | AAW23237 GCAP-II C |
| 19         | 375   | 64.8        | 70     | 2     | AAW18471 Human GCA |
| 20         | 375   | 64.8        | 70     | 2     | AAW18480 Human GCA |
| 21         | 375   | 64.8        | 70     | 2     | AAW23225 GCAP-II C |
| 22         | 370   | 63.9        | 69     | 2     | AAW18472 Human GCA |
| 23         | 370   | 63.9        | 69     | 2     | AAW18481 Human GCA |
| 24         | 370   | 63.9        | 69     | 2     | AAW23226 GCAP-II C |
| 25         | 362   | 62.5        | 67     | 2     | AAW18474 Human GCA |

ALIGNMENTS

RESULT 1

|          |    |       |      |     |   |          |           |
|----------|----|-------|------|-----|---|----------|-----------|
| AAW18498 | 26 | 362   | 62.5 | 67  | 2 | AAW23228 | GCAP-II C |
| XX       | 27 | 360   | 62.2 | 69  | 2 | AAW18488 | Human GCA |
| XX       | 28 | 356   | 61.5 | 66  | 2 | AAW18491 | Human GCA |
| XX       | 29 | 356   | 61.5 | 66  | 2 | AAW23238 | GCAP-II C |
| XX       | 30 | 356   | 61.5 | 106 | 8 | ADR45874 | Amino aci |
| XX       | 31 | 347   | 59.9 | 64  | 2 | AAW18492 | Human GCA |
| XX       | 32 | 347   | 59.9 | 64  | 2 | AAW23239 | GCAP-II C |
| XX       | 33 | 300   | 51.8 | 56  | 2 | AAW18469 | Human GCA |
| XX       | 34 | 300   | 51.8 | 56  | 2 | AAW23223 | GCAP-II C |
| XX       | 35 | 252   | 43.5 | 109 | 6 | ABG74824 | Oposum 1  |
| XX       | 36 | 230   | 39.7 | 43  | 2 | AAW18489 | Human GCA |
| XX       | 37 | 230   | 39.7 | 43  | 2 | AAW23236 | GCAP-II C |
| XX       | 38 | 216   | 37.3 | 85  | 3 | AAW53987 | Human col |
| XX       | 39 | 203   | 35.1 | 38  | 2 | AAW18475 | Human GCA |
| XX       | 40 | 203   | 35.1 | 38  | 2 | AAW23229 | GCAP-II C |
| XX       | 41 | 197   | 34.0 | 37  | 2 | AAW18493 | Human GCA |
| XX       | 42 | 197   | 34.0 | 37  | 2 | AAW23240 | GCAP-II C |
| XX       | 43 | 164.5 | 28.4 | 115 | 7 | ADJ18912 | Human NOV |
| XX       | 44 | 162.5 | 28.1 | 94  | 2 | AAW62442 | Guanylate |
| XX       | 45 | 162.5 | 28.1 | 94  | 2 | AAW69810 | Expressio |

AAW18498 standard; protein; 112 AA.

AC AAW18498;

DT 22-APR-1998 (first entry)

DE Human GCAP-II precursor protein.

XX Guanyl cyclase C activating peptide II; GCAP-II; insulinotropic; diabetes; endocrine disorder; diagnosis; treatment; human.

OS Homo sapiens.

PN DE19543628-A1.

XX

PD 28-MAY-1997.

XX

PF 24-NOV-1995; 95DE-01043628.

XX

PR 24-NOV-1995; 95DE-01043628.

XX

PA (FORS/) FORSMANN W.

XX

PI Forssmann W, Kist A, Kruhoffer M, Meyer M, Pardigol A, Heine G;

XX

DR WPI; 1997-290350/27.

XX

DR N-PSDB; AAT65115.

XX

PT New guanyl cyclase C activating peptide fragments - have insulinotropic activity, useful for treating diabetes, etc.

XX

PS Example 6; Fig 11; 33pp; German.

XX

CC This sequence represents a precursor of the guanyl cyclase C activating peptide, GCAP-II, which affects insulin secretion by the beta cells in the pancreas. This peptide is useful for treating pancreatic endocrine disorders, especially diabetes mellitus type II, renal and intestinal disorders, disorders of the gastrointestinal, respiratory and urogenital apparatus, disorders of the cardiovascular and nervous systems, disorders of the integuments and sense organs and diseases associated with GCAP-II (89-112) deficiency. This peptide can be used for treatment of electrolyte effects on bone reconstruction (osteoporosis) or the dental apparatus. Antibodies to GCAP-II (89-112) can be used to treat diseases associated with overproduction of GCAP-II (89-112). Human GCAP-II (89-112) and GCAP-I (99-15) cDNA are useful for diagnosis and treatment of the above disorders e.g. gene therapy for diabetes

```
XX SQ Sequence 112 AA;
Query Match 100.0%; Score 579; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 4.5e-57;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGCRAASGLLPVAVVLLLLQSTQSVYIQGFRVQLSMKLSDLAQWAPSRLQAO 60
DB 1 MGCRAASGLLPVAVVLLLLQSTQSVYIQGFRVQLSMKLSDLAQWAPSRLQAO 60
QY 61 SLLPAVCHHPALPQDLQPVCAQSEASSIFKTLRTIANDDCCLCVNACTGCL 112
DB 61 SLLPAVCHHPALPQDLQPVCAQSEASSIFKTLRTIANDDCCLCVNACTGCL 112

RESULT 2
AAW10595
ID AAW10595 standard; protein; 112 AA.
XX AAW10595;
AC AAW10595;
DT 29-OCT-1997 (first entry)
XX
DE Guanylate cyclase activating peptide II.
XX
XX Human; guanylate cyclase; activating peptide; GCAP-II; cGMP;
KW transepithelial transport; treatment; kidney; intestinal; respiratory;
KW urogenital; circulatory; nervous system; disorder; disease; endocrine;
KW sensory; system; osteoporosis; dental; pancreas; diabetes; hypophysis;
KW gastrointestinal tract; diarrhoea; gene therapy; probe;
KW recombinant production; transgenic animal; antibody; immunoassay reagent.
XX
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..88
FT /label= sig_peptide
FT 89..112
FT /label= mat_peptide
XX
PN DE19528544-A1.
XX
XX 06-FEB-1997.
XX
XX 03-AUG-1995; 95DE-01028544.
XX
XX 03-AUG-1995; 95DE-01028544.
XX (FORS/) FORSSMANN W.
XX
XX Forssmann W;
XX
XX WPI; 1997-110032/11.
XX N-PSDB; AAT60819.
XX
XX Guanylate cyclase activating peptide II - increases cGMP formation, and
XX controls transport of water and electrolytes across epithelial cells.
XX
XX Claim 1; Page 4; 15pp; German.
XX
XX The present sequence is the human guanylate cyclase activating peptide II
XX (GCAP-II), which increases cGMP formation, and is involved in the control
XX of transepithelial water and electrolyte transport. GCAP-II can be used
XX to treat a variety of kidney, intestinal, respiratory, urogenital,
XX circulatory and nervous system disorders, diseases of the endocrine and
XX sensory systems (e.g. osteoporosis, and dental disease), disorders of the
XX pancreas (e.g. diabetes, and hypophysis) or the endocrine
XX gastrointestinal tract and for the long term treatment of diarrhoea,
XX without inducing an immune response. The GCAP-II cDNA can be used to
XX treat the same conditions, clone the GCAP-II-encoding gene for use in
XX gene therapy, as a hybridisation probe and for the production of
XX recombinant GCAP-II or transgenic animal creation. Antibodies raised
```

```
CC against GCAP-II are useful as immunoassay reagents. GCAP-II is
CC administered at, e.g. 100-1200 microg/day by intravenous or intramuscular
CC injection or 300-1200 microg/day subcutaneously. It may also be given
CC orally, intranasally or by inhalation, in typical unit doses of 0.3-30
CC mg. GCAP-II was chemically synthesised, or isolated by chromatography
CC from transformed eukaryotic or prokaryotic cells, or human blood. When
CC T84 cells were incubated with synthetic GCAP-II, generation of cGMP was
CC increased in a dose dependent manner. GCAP-II influences cGMP production
CC via a known receptor for heat stable enterotoxin. Other stomach,
CC intestinal, pancreatic and liver cells also responded to GCAP-II, e.g.
CC via changes in intracellular Ca2+ ion concentration
XX
XX Sequence 112 AA;
```

```
Query Match 100.0%; Score 579; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 4.5e-57;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGCRAASGLLPVAVVLLLLQSTQSVYIQGFRVQLSMKLSDLAQWAPSRLQAO 60
DB 1 MGCRAASGLLPVAVVLLLLQSTQSVYIQGFRVQLSMKLSDLAQWAPSRLQAO 60
QY 61 SLLPAVCHHPALPQDLQPVCAQSEASSIFKTLRTIANDDCCLCVNACTGCL 112
DB 61 SLLPAVCHHPALPQDLQPVCAQSEASSIFKTLRTIANDDCCLCVNACTGCL 112
```

```
RESULT 3
ASG74823
ID ABG74823 standard; protein; 112 AA.
XX
XX AC ABG74823;
XX
XX DT 12-JUN-2003 (first entry)
XX
XX Human uroguanylin precursor SEQ ID 5.
XX
XX Apical membrane; mucosal epithelial cell; respiratory tract;
KW guanylate cyclase C; G protein-coupled receptor; guanosine triphosphate;
KW cyclic guanosine monophosphate; cGMP; chloride ion secretion; inhalation;
KW membrane-associated type II protein kinase; mucus fluidisation;
KW cystic fibrosis transmembrane conductance regulator; breathing disorder;
KW mucus secretion; antiasthmatic; antiinflammatory; bronchial asthma;
KW chronic bronchitis; cystic fibrosis; uroguanylin; human.
XX
XX Homo sapiens.
XX
XX PN WO200298912-A2.
XX
XX PD 12-DEC-2002.
XX
XX 05-JUN-2002; 2002WO-DE002040.
XX
XX 05-JUN-2001; 2001DE-01027119.
XX
XX (CETI/) CETIN Y.
XX (SAVA/) SAVAS Y.
XX
XX Cetin Y, Savas Y;
XX
XX WPI; 2003-156842/15.
```

```
XX Composition useful for treating respiratory disease, comprises a peptide
XX that activates guanylate cyclase C, and is delivered to the apical
XX membrane through the respiratory tract.
XX
XX Claim 3; Page 3; 23pp; German.
XX
XX This invention describes a novel medicament in a formulation that is
XX delivered to the apical membrane of mucosal epithelial cells through the
XX respiratory tract. The medicament contains at least one peptide that
XX activates guanylate cyclase C (GCC). GCC is a G protein-coupled receptor
XX that catalyses conversion of guanosine triphosphate to cyclic guanosine
```

CC monophosphate (cGMP) and is present on the apical (air) side of  
 CC respiratory epithelial cells but not on the basolateral (blood) side.  
 CC cGMP activates membrane-associated type II protein kinase which in turn  
 CC activates the regulatory domain of the cystic fibrosis transmembrane  
 CC conductance regulator, resulting in secretion of chloride ions and water  
 CC from the cells, causing fluidisation of the mucus. The products of the  
 CC invention are used to make an inhalation device containing the medicament  
 CC for diagnosing diseases that are accompanied by breathing disorders or  
 CC disorders of mucus secretion in the respiratory tract, by detecting at  
 CC least one GCC activator. The products of the invention have antiasthmatic  
 CC and antiinflammatory activity. The method is useful for diagnosing and  
 CC treating diseases accompanied by breathing disorders or disorders of  
 CC mucus secretion in the respiratory tract particularly bronchial asthma,  
 CC chronic bronchitis and cystic fibrosis. The product of the invention  
 CC improves fluidity and evacuation of bronchial mucus and acts locally  
 CC (since the medicament does not enter the bloodstream), so systemic side  
 CC effects are minimised. Only very small doses of the medicament are  
 CC required. This sequence represents a precursor of the human guanylate  
 CC cyclase C activator peptide, uroguanylin, described in the disclosure of  
 CC the invention

XX SQ Sequence 112 AA;

Query Match 100.0%; Score 579; DB 6; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-57;  
 Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGCRAASGLPGVAVVLLLLQLSTQSVYIQGFRVQLSKKLSDLQAOWAPSPRLQAQ 60  
 Db 1 MGCRAASGLPGVAVVLLLLQLSTQSVYIQGFRVQLSKKLSDLQAOWAPSPRLQAQ 60  
 Qy 61 SLLPAVCHHPALPDQLPVCASQEAASSIFKTLRTIANDDCELCVNVACTGCL 112  
 Db 61 SLLPAVCHHPALPDQLPVCASQEAASSIFKTLRTIANDDCELCVNVACTGCL 112

# RESULT 4

ADRA45875  
 ID ADRA45875 standard; peptide; 112 AA.

AC ADRA45875;

DT 18-NOV-2004 (first entry)

XX Amino acid sequence of ST receptor ligand uroguanylin.

XX heat stable toxin; ST; ST receptor; colorectal cancer; gastric cancer;  
 KW oesophageal cancer; colorectal cell; gastric cell; oesophageal cell;  
 KW uroguanylin.

XX Homo sapiens.

XX WO2004071436-A2.

XX 26-AUG-2004.

XX 10-FEB-2004; 2004WO-US003765.

XX 10-FEB-2003; 2003US-0446730P.

XX (UYUE-) UNIV JEFFERSON THOMAS.

XX Waldman SA, Pitari GM, Park J, Schulz S, Wolfe HR;

XX WPI; 2004-615913/59.

XX Increasing ST receptor molecules on the surface of a colorectal, gastric  
 PT or oesophageal cancer cell to treat these cancers comprises administering  
 PT ST receptor ligand molecules that bind to ST receptors on the surface of  
 PT the cancer cell.

XX Claim 6; SEQ ID NO 56; 97pp; English.

XX

CC The specification describes a method for increasing the number of heat  
 CC stable toxin (ST) receptor molecules on the surface of a metastasised  
 CC colorectal cancer cell. The method comprises administering, by continuous  
 CC infusion, at least 0.1 nM of an ST receptor ligand per 10 kg bodyweight  
 CC per hour for at least 6 hours, where ST receptor ligand molecules bind to  
 CC ST receptors on the surface of the cancer cell in the individual and the  
 CC number of ST receptor molecules on the surface of the cancer cell is  
 CC increased. Therapeutic compositions comprising components which target ST  
 CC receptors can then be used to inhibit proliferation of the colorectal,  
 CC gastric and oesophageal cells. This method may be used for treating  
 CC individuals that have diseases that affect colorectal, gastric and  
 CC oesophageal cells, including colorectal, gastric or oesophageal cancers.  
 CC The present sequence represents a ST receptor ligand used in the method  
 CC of the invention, designated uroguanylin.

XX SQ Sequence 112 AA;

Query Match 100.0%; Score 579; DB 8; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-57;  
 Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGCRAASGLPGVAVVLLLLQLSTQSVYIQGFRVQLSKKLSDLQAOWAPSPRLQAQ 60

Db 1 MGCRAASGLPGVAVVLLLLQLSTQSVYIQGFRVQLSKKLSDLQAOWAPSPRLQAQ 60

Qy 61 SLLPAVCHHPALPDQLPVCASQEAASSIFKTLRTIANDDCELCVNVACTGCL 112

Db 61 SLLPAVCHHPALPDQLPVCASQEAASSIFKTLRTIANDDCELCVNVACTGCL 112

# RESULT 5

AAW18467  
 ID AAW18467 standard; peptide; 108 AA.

AC AAW18467;

XX 23-APR-1998 (first entry)

XX Human GCAP-II (89-112) endoprotease Arg-C digested fragment 1.

XX Guanyl cyclase C activating peptide II; GCAP-II; insulinotropic;  
 KW diabetes; endocrine disorder; diagnosis; treatment; human.

XX Homo sapiens.

XX DE19543628-A1.

XX 28-MAY-1997.

XX 24-NOV-1995; 95DE-01043628.

XX 24-NOV-1995; 95DE-01043628.

XX (FORS/) FORSMANN W.

XX Forssmann W, Kist A, Kruhoffer M, Meyer M, Pardigol A, Heine G;

XX WPI; 1997-290350/27.

XX New guanyl cyclase C activating peptide fragments - have insulinotropic  
 PT activity, useful for treating diabetes, etc.

XX Claim 3; Fig 3; 33pp; German.

XX Peptides AAW18467-W18470 represent fragments of the guanyl cyclase C  
 CC activating peptide, GCAP-II, obtained by digestion with endoprotease Arg-  
 CC C. GCAP-II is involved in insulin secretion by pancreatic beta cells.  
 CC This peptide fragment could be used to which affects insulin secretion by  
 CC the beta cells treat pancreatic endocrine disorders, especially diabetes  
 CC mellitus type II, renal and intestinal disorders, disorders of the  
 CC gastrointestinal, respiratory and urogenital apparatus, disorders of the  
 CC cardiovascular and nervous systems, disorders of the integuments and  
 CC sense organs and diseases associated with GCAP II (89-112) deficiency.



CC increased in a dose dependent manner. GCAP-II influences cGMP production  
CC via a known receptor for heat stable enterotoxin. Other stomach,  
CC intestinal, pancreatic and liver cells also responded to GCAP-II, e.g.  
CC via changes in intracellular Ca<sup>2+</sup> ion concentration  
XX  
SQ Sequence 108 AA;

Query Match 95.7%; Score 554; DB 2; Length 108;  
Best Local Similarity 100.0%; Pred. No. 2.8e-54;  
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AASGLPGVAVVLLLLQSTQSVVYIQGFRVQLESMMKLSDLAQWAPSRLQAOQLLP 64  
Db 1 AASGLPGVAVVLLLLQSTQSVVYIQGFRVQLESMMKLSDLAQWAPSRLQAOQLLP 60  
Qy 65 AVCHHPALPQDLQPVCAQSEASSIFKTLRTIANDDCELCVNVACTGCL 112  
Db 61 AVCHHPALPQDLQPVCAQSEASSIFKTLRTIANDDCELCVNVACTGCL 108

RESULT 8  
AAW18484  
ID AAW18484 standard; peptide; 84 AA.

AC AAW18484;

XX 22-APR-1998 (first entry)

XX Human GCAP-II (89-112) chymotrypsin digested fragment 1.

XX Guanyl cyclase C activating peptide II; GCAP-II; insulinotropic;  
KW diabetes; endocrine disorder; diagnosis; treatment; human.

XX Homo sapiens.

XX DE19543628-AL.

XX 28-MAY-1997.

XX 24-NOV-1995; 95DE-01043628.

XX 24-NOV-1995; 95DE-01043628.

XX (FORS/) FORSMANN W.

XX Forsemann W, Kist A, Kruhoeffer M, Meyer M, Pardigol A, Heine G;

XX WPI; 1997-290350/27.

XX New guanyl cyclase C activating peptide fragments - have insulinotropic  
PT activity, useful for treating diabetes, etc.

XX Claim 3; Fig 3; 33pp; German.

XX Peptides AAW18484-W18487 represent fragments of the guanyl cyclase C  
CC activating peptide, GCAP-II, obtained by digestion with chymotrypsin.  
CC GCAP-II is involved in insulin secretion by pancreatic beta cells. This  
CC peptide fragment could be used to which affects insulin secretion by the  
CC beta cells treat pancreatic endocrine disorders, especially diabetes  
CC mellitus type II, renal and intestinal disorders, disorders of the  
CC respiratory, gastrointestinal and urogenital apparatus, disorders of the  
CC cardiovascular and nervous systems, disorders of the integuments and  
CC sense organs and diseases associated with GCAP II (89-112) deficiency.  
CC This peptide can be used for treatment of electrolyte effects on bone  
CC reconstruction (osteoporosis) or the dental apparatus. Antibodies to GCAP  
CC -II (89-112) can be used to treat diseases associated with overproduction  
CC of GCAP-II (89-112). Human GCAP-II (89-112) and GCAP I (99-15) cDNA are  
CC useful for diagnosis and treatment of the above disorders e.g. gene  
CC therapy for diabetes  
XX

SQ Sequence 84 AA;

Query Match 76.9%; Score 445; DB 2; Length 84;

Best Local Similarity 100.0%; Pred. No. 3.9e-42;  
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 IOYQGFVQLESMMKLSDLAQWAPSRLQAOQLLPVCHHPALPQDLQPVCAQSEASSI 88  
Db 1 IOYQGFVQLESMMKLSDLAQWAPSRLQAOQLLPVCHHPALPQDLQPVCAQSEASSI 60

Qy 89 FXTLRTIANDDCELCVNVACTGCL 112  
Db 61 FXTLRTIANDDCELCVNVACTGCL 84

RESULT 9  
AAW23232  
ID AAW23232 standard; peptide; 84 AA.

XX AAW23232;

XX 29-OCT-1997 (first entry)

XX GCAP-II C-terminal fragment prepared by chymotrypsin.

XX Human; guanylate cyclase; activating peptide; GCAP-II; cGMP;  
KW transepithelial transport; treatment; kidney; intestinal; respiratory;  
KW urogenital; circulatory; nervous system; disorder; disease; endocrine;  
KW sensory; system; osteoporosis; dental; pancreas; diabetes; hypophysis;  
KW gastrointestinal tract; diarrhoea; gene therapy; probe;  
KW recombinant production; transgenic animal; antibody; immunoassay reagent.

XX Homo sapiens.

XX DE19528544-AL.

XX 06-FEB-1997.

XX 03-AUG-1995; 95DE-01028544.

XX 03-AUG-1995; 95DE-01028544.

XX (FORS/) FORSMANN W.

XX Forsemann W;

XX WPI; 1997-110032/11.

XX Guanylate cyclase activating peptide II - increases cGMP formation, and  
PT controls transport of water and electrolytes across epithelial cells.

XX Claim 3; Page 6; 15pp; German.

XX The present sequence is a carboxy-terminal fragment of the human  
CC guanylate cyclase activating peptide II (GCAP-II) precursor, prepared by  
CC endoproteolytic cleavage with chymotrypsin. GCAP-II increases cGMP  
CC formation, and is involved in the control of transepithelial water and  
CC electrolyte transport. GCAP-II can be used to treat a variety of kidney,  
CC intestinal, respiratory, urogenital, circulatory and nervous system  
CC disorders, diseases of the endocrine and sensory systems (e.g.  
CC osteoporosis, and dental disease), disorders of the pancreas (e.g.  
CC diabetes, and hypophysis) or the endocrine gastrointestinal tract and for  
CC the long term treatment of diarrhoea, without inducing an immune  
CC response. The GCAP-II cDNA can be used to treat the same conditions,  
CC clone the GCAP-II-encoding gene for use in gene therapy, as a  
CC hybridisation probe and for the production of recombinant GCAP-II or  
CC transgenic animal creation. Antibodies raised against GCAP-II are useful  
CC as immunoassay reagents. GCAP-II, or a fragment, are administered at,  
CC e.g. 100-1200 microg/day by intravenous or intramuscular injection or 300  
CC -1200 microg/day subcutaneously. They may also be given orally,  
CC intranasally or by inhalation, in typical unit doses of 0.3-30 mg. GCAP-  
CC II was chemically synthesised, or isolated by chromatography from  
CC transformed eukaryotic or prokaryotic cells, or human blood. When T84  
CC cells were incubated with synthetic GCAP-II, generation of cGMP was  
CC increased in a dose dependent manner. GCAP-II influences cGMP production  
CC via a known receptor for heat stable enterotoxin. Other stomach,

CC intestinal, pancreatic and liver cells also responded to GCAP-II, e.g.  
 CC via changes in intracellular Ca2+ ion concentration

SQ Sequence 84 AA;

Query Match 76.9%; Score 445; DB 2; Length 84;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-42;  
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 IQQFRVQLESMMKLSLEAQWAPSPRLQAQSLLPVCHHPALPDQLQPVCAEQEASSI 88  
 DB 1 IQQFRVQLESMMKLSLEAQWAPSPRLQAQSLLPVCHHPALPDQLQPVCAEQEASSI 60  
 QY 89 FKTLRTIANDDCELCVNVACTGCL 112  
 DB 61 FKTLRTIANDDCELCVNVACTGCL 84

## RESULT 10

AAW18485  
 ID AAW18485 standard; peptide; 81 AA.

AC AAW18485;

DT 22-APR-1998 (first entry)

XX Human GCAP-II (89-112) chymotrypsin digested fragment 2.

XX Guanyl cyclase C activating peptide II; GCAP-II; insulinotropic;  
 KW diabetes; endocrine disorder; diagnosis; treatment; human.

XX Homo sapiens.

XX DE19543628-A1.

XX 28-MAY-1997.

XX 24-NOV-1995; 95DE-01043628.

XX 24-NOV-1995; 95DE-01043628.

XX (FORS/) FORSSMANN W.

XX Forssmann W, Kist A, Kruhoffer M, Meyer M, Pardigol A, Heine G;  
 WPI; 1997-290350/27.

XX New guanyl cyclase C activating peptide fragments - have insulinotropic  
 activity, useful for treating diabetes, etc.

XX Claim 3; Fig 3; 33pp; German.

XX Peptides AAW18484-W18487 represent fragments of the guanyl cyclase C  
 activating peptide, GCAP-II, obtained by digestion with chymotrypsin.  
 CC GCAP-II is involved in insulin secretion by pancreatic beta cells. This  
 CC peptide fragment could be used to which affects insulin secretion by the  
 CC beta cells treat pancreatic endocrine disorders, especially diabetes  
 CC mellitus type II, renal and intestinal disorders, disorders of the  
 CC respiratory, gastrointestinal and urogenital apparatus, disorders of the  
 CC cardiovascular and nervous systems, disorders of the integuments and  
 CC sense organs and diseases associated with GCAP II (89-112) deficiency.  
 CC This peptide can be used for treatment of electrolyte effects on bone  
 CC reconstruction (osteoporosis) or the dental apparatus. Antibodies to GCAP  
 CC -II (89-112) can be used to treat diseases associated with overproduction  
 CC of GCAP-II (89-112). Human GCAP-II (89-112) and GCAP I (99-15) cDNA are  
 CC useful for diagnosis and treatment of the above disorders e.g. gene  
 CC therapy for diabetes

XX Sequence 81 AA;

Query Match 74.1%; Score 429; DB 2; Length 81;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-40;  
 Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 QGFRVQLESMMKLSLEAQWAPSPRLQAQSLLPVCHHPALPDQLQPVCAEQEASSI 91  
 DB 1 QGFRVQLESMMKLSLEAQWAPSPRLQAQSLLPVCHHPALPDQLQPVCAEQEASSI 60  
 QY 92 LRTIANDDCELCVNVACTGCL 112  
 DB 61 LRTIANDDCELCVNVACTGCL 81

## RESULT 11

AAW23233  
 ID AAW23233 standard; peptide; 81 AA.

AC AAW23233;

DT 29-OCT-1997 (first entry)

XX GCAP-II C-terminal fragment prepared by chymotrypsin.

XX Human; guanylate cyclase; activating peptide; GCAP-II; cGMP;  
 KW transepithelial transport; treatment; kidney; intestinal; respiratory;  
 KW urogenital; circulatory; nervous system; disorder; disease; endocrine;  
 KW sensory; system; osteoporosis; dental; pancreas; diabetes; hypophysis;  
 KW gastrointestinal tract; diarrhoea; gene therapy; probe;  
 KW recombinant production; transgenic animal; antibody; immunoassay reagent.

XX Homo sapiens.

XX DE19528544-A1.

XX 06-FEB-1997.

XX 03-AUG-1995; 95DE-01028544.

XX 03-AUG-1995; 95DE-01028544.

XX (FORS/) FORSSMANN W.

XX Forssmann W;

XX WPI; 1997-110032/11.

XX Guanylate cyclase activating peptide II - increases cGMP formation, and  
 controls transport of water and electrolytes across epithelial cells.

XX Claim 3; Page 6; 15pp; German.

XX The present sequence is a carboxy-terminal fragment of the human  
 CC guanylate cyclase activating peptide II (GCAP-II) precursor, prepared by  
 CC endoproteolytic cleavage with chymotrypsin. GCAP-II increases cGMP  
 CC formation, and is involved in the control of transepithelial water and  
 CC electrolyte transport. GCAP-II can be used to treat a variety of kidney,  
 CC intestinal, respiratory, urogenital, circulatory and nervous system  
 CC disorders, diseases of the endocrine and sensory systems (e.g.  
 CC osteoporosis, and dental disease), disorders of the pancreas (e.g.  
 CC diabetes, and hypophysis) or the endocrine gastrointestinal tract and for  
 CC the long term treatment of diarrhoea, without inducing an immune  
 CC response. The GCAP-II cDNA can be used to treat the same conditions,  
 CC clone the GCAP-II-encoding gene for use in gene therapy, as a  
 CC hybridisation probe and for the production of recombinant GCAP-II or  
 CC transgenic animal creation. Antibodies raised against GCAP-II are useful  
 CC as immunoassay reagents. GCAP-II, or a fragment, are administered at,  
 CC e.g. 100-1200 microg/day by intravenous or intramuscular injection or 300  
 CC -1200 microg/day subcutaneously. They may also be given orally.  
 CC intranasally or by inhalation, in typical unit doses of 0.3-30 mg. GCAP-  
 CC II was chemically synthesised, or isolated by chromatography from T84  
 CC transformed eukaryotic or prokaryotic cells, or human blood. When T84  
 CC cells were incubated with synthetic GCAP-II, generation of cGMP was  
 CC increased in a dose dependent manner. GCAP-II influences cGMP production  
 CC via a known receptor for heat stable enterotoxin. Other stomach,  
 CC intestinal, pancreatic and liver cells also responded to GCAP-II, e.g.  
 CC via changes in intracellular Ca2+ ion concentration



XX Sequence 81 AA;  
SQ

Query Match 74.1%; Score 429; DB 2; Length 81;  
Best Local Similarity 100.0%; Pred. No. 2.4e-40;  
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 QGFRVQLESMMKSLDLEAQWAPSPRLQASLLPAVCHHPALPDQLQPVCAEQEASSIFKTLRT 91  
Db 1 QGFRVQLESMMKSLDLEAQWAPSPRLQASLLPAVCHHPALPDQLQPVCAEQEASSIFKTLRT 60

Qy 92 LRTIANDDCELCVNVACTGCL 112  
Db 61 LRTIANDDCELCVNVACTGCL 81

RESULT 12  
AAW18486  
ID AAW18486 standard; peptide; 78 AA.  
XX  
AC AAW18486;  
XX  
DT 22-APR-1998 (first entry)  
XX  
DE Human GCAP-II (89-112) chymotrypsin digested fragment 3.  
XX  
KW Guanyl cyclase C activating peptide II; GCAP-II; insulinotropic;  
KW diabetes; endocrine disorder; diagnosis; treatment; human.  
XX  
OS Homo sapiens.  
XX  
PN DE19543628-AL.  
XX  
PD 28-MAY-1997.  
XX  
PF 24-NOV-1995; 95DE-01043628.  
XX  
PR 24-NOV-1995; 95DE-01043628.  
XX  
PA (FORS/) FORSMANN W.  
XX  
PI Forsemann W, Kist A, Krühoffer M, Meyer M, Pardigol A, Heine G;  
XX WPI; 1997-290350/27.  
XX  
PT New guanyl cyclase C activating peptide fragments - have insulinotropic  
PT activity, useful for treating diabetes, etc.  
XX  
PS Claim 3; Fig 3; 33pp; German.  
XX

Peptides AAW18484-W18487 represent fragments of the guanyl cyclase C activating peptide, GCAP-II, obtained by digestion with chymotrypsin. GCAP-II is involved in insulin secretion by pancreatic beta cells. This peptide fragment could be used to which affects insulin secretion by the beta cells treat pancreatic endocrine disorders, especially diabetes mellitus type II, renal and intestinal disorders, disorders of the respiratory, gastrointestinal and urogenital apparatus, disorders of the cardiovascular and nervous systems, disorders of the integuments and sense organs and diseases associated with GCAP II (89-112) deficiency. This peptide can be used for treatment of electrolyte effects on bone reconstruction (osteoporosis) or the dental apparatus. Antibodies to GCAP -II (89-112) can be used to treat diseases associated with overproduction of GCAP-II (89-112). Human GCAP-II (89-112) and GCAP I (99-15) cDNA are useful for diagnosis and treatment of the above disorders e.g. gene therapy for diabetes

XX  
SQ Sequence 78 AA;  
Query Match 71.2%; Score 412; DB 2; Length 78;  
Best Local Similarity 100.0%; Pred. No. 1.9e-38;  
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 RVQLESMMKSLDLEAQWAPSPRLQASLLPAVCHHPALPDQLQPVCAEQEASSIFKTLRT 94

Db 1 RVQLESMMKSLDLEAQWAPSPRLQASLLPAVCHHPALPDQLQPVCAEQEASSIFKTLRT 60  
Qy 95 IANDDCELCVNVACTGCL 112  
Db 61 IANDDCELCVNVACTGCL 78

RESULT 13  
AAW23234  
ID AAW23234 standard; peptide; 78 AA.  
XX  
AC AAW23234;  
XX  
DT 29-OCT-1997 (first entry)  
XX  
DE GCAP-II C-terminal fragment prepared by chymotrypsin.  
XX  
KW Human; guanylate cyclase; activating peptide; GCAP-II; cGMP;  
KW transepithelial transport; treatment; kidney; intestinal; respiratory;  
KW urogenital; circulatory; nervous system; disorder; disease; endocrine;  
KW sensory; system; osteoporosis; dental; pancreas; diabetes; hypophysis;  
KW gastrointestinal tract; diarrhoea; gene therapy; probe;  
KW recombinant production; transgenic animal; antibody; immunoassay reagent.  
XX  
OS Homo sapiens.  
XX  
PN DE19528544-AL.  
XX  
PD 06-FEB-1997.  
XX  
PF 03-AUG-1995; 95DE-01028544.  
XX  
PR 03-AUG-1995; 95DE-01028544.  
XX  
PA (FORS/) FORSMANN W.  
XX  
PI Forsemann W;  
XX  
DR WPI; 1997-110032/11.  
XX

Guanylate cyclase activating peptide II - increases cGMP formation, and controls transport of water and electrolytes across epithelial cells.

Claim 3; Page 6; 15pp; German.

The present sequence is a carboxy-terminal fragment of the human guanylate cyclase activating peptide II (GCAP-II) precursor, prepared by endoproteolytic cleavage with chymotrypsin. GCAP-II increases cGMP formation, and is involved in the control of transepithelial water and electrolyte transport. GCAP-II can be used to treat a variety of kidney, intestinal, respiratory, urogenital, circulatory and nervous system disorders, diseases of the endocrine and sensory systems (e.g. osteoporosis, and dental disease), disorders of the pancreas (e.g. diabetes, and hypophysis) or the endocrine gastrointestinal tract and for the long term treatment of diarrhoea, without inducing an immune response. The GCAP-II cDNA can be used to treat the same conditions, clone the GCAP-II-encoding gene for use in gene therapy, as a hybridisation probe and for the production of recombinant GCAP-II or transgenic animal creation. Antibodies raised against GCAP-II are useful as immunoassay reagents. GCAP-II, or a fragment, are administered at, e.g. 100-1200 microg/day by intravenous or intramuscular injection or 300 -1200 microg/day subcutaneously. They may also be given orally, intranasally or by inhalation, in typical unit doses of 0.3-30 mg. GCAP-II was chemically synthesised, or isolated by chromatography from transformed eukaryotic or prokaryotic cells, or human blood. When T84 cells were incubated with synthetic GCAP-II, generation of cGMP was increased in a dose dependent manner. GCAP-II influences cGMP production via a known receptor for heat stable enterotoxin. Other stomach, intestinal, pancreatic and liver cells also responded to GCAP-II, e.g. via changes in intracellular Ca<sup>2+</sup> ion concentration

XX  
SQ Sequence 78 AA;

```
Query Match 71.2%; Score 412; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 1.9e-38;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 RVQESMKKLSLEAQWAPSPRLQASLLPAVCHHPALPDQLQPVCAQSEASSIFKTLRT 94
DB 1 RVQESMKKLSLEAQWAPSPRLQASLLPAVCHHPALPDQLQPVCAQSEASSIFKTLRT 60

QY 95 IANDCELVCNVACTGCL 112
DB 61 IANDCELVCNVACTGCL 78

RESULT 14
AAW18468
ID AAW18468 standard; peptide; 77 AA.
AC AAW18468;
XX
XX 23-APR-1998 (first entry)
XX
XX Human GCAP-II (89-112) endoprotease Arg-C digested fragment 2.
XX
XX Guanyl cyclase C activating peptide II; GCAP-II; insulinotropic;
XX diabetes; endocrine disorder; diagnosis; treatment; human.
XX
XX Homo sapiens.
XX
XX DE19543628-A1.
XX
XX 28-MAY-1997.
XX
XX 24-NOV-1995; 95DE-01043628.
XX
XX 24-NOV-1995; 95DE-01043628.
XX
XX (FORS/) FORSSMANN W.
XX
XX Forssmann W, Kist A, Kruhoeffner M, Meyer M, Pardigol A, Heine G;
XX WPI; 1997-290350/27.
XX
XX New guanyl cyclase C activating peptide fragments - have insulinotropic
XX activity, useful for treating diabetes, etc.
XX
XX Claim 3; Fig 3; 33pp; German.
XX
XX Peptides AAW18467-W18470 represent fragments of the guanyl cyclase C
XX activating peptide, GCAP-II, obtained by digestion with endoprotease Arg-
XX C. GCAP-II is involved in insulin secretion by pancreatic beta cells.
XX This peptide fragment could be used to which affects insulin secretion by
XX the beta cells treat pancreatic endocrine disorders, especially diabetes
XX mellitus type II, renal and intestinal disorders, disorders of the
XX gastrointestinal, respiratory and urogenital apparatus, disorders of the
XX sense organs and diseases associated with GCAP II (89-112) deficiency.
XX This peptide can be used for treatment of electrolyte effects on bone
XX reconstruction (osteoporosis) or the dental apparatus. Antibodies to GCAP
XX -II (89-112) can be used to treat diseases associated with overproduction
XX of GCAP-II (89-112). Human GCAP-II (89-112) and GCAP I (99-15) cDNA are
XX useful for diagnosis and treatment of the above disorders e.g. gene
XX therapy for diabetes
XX
XX Sequence 77 AA;
XX
XX Query Match 70.3%; Score 407; DB 2; Length 77;
XX Best Local Similarity 100.0%; Pred. No. 6.7e-38;
XX Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 VQESMKKLSLEAQWAPSPRLQASLLPAVCHHPALPDQLQPVCAQSEASSIFKTLRTI 95
DB 1 VQESMKKLSLEAQWAPSPRLQASLLPAVCHHPALPDQLQPVCAQSEASSIFKTLRTI 60

Query Match 70.3%; Score 407; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 6.7e-38;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 VQESMKKLSLEAQWAPSPRLQASLLPAVCHHPALPDQLQPVCAQSEASSIFKTLRTI 95
DB 1 VQESMKKLSLEAQWAPSPRLQASLLPAVCHHPALPDQLQPVCAQSEASSIFKTLRTI 60
```

```
QY 96 ANDDCELVCNVACTGCL 112
DB 61 ANDDCELVCNVACTGCL 77

RESULT 15
AAW18479
ID AAW18479 standard; peptide; 77 AA.
XX
XX AAW18479;
XX
XX 22-APR-1998 (first entry)
XX
XX Human GCAP-II (89-112) trypsin digested fragment 2.
XX
XX Guanyl cyclase C activating peptide II; GCAP-II; insulinotropic;
XX diabetes; endocrine disorder; diagnosis; treatment; human.
XX
XX Homo sapiens.
XX
XX DE19543628-A1.
XX
XX 28-MAY-1997.
XX
XX 24-NOV-1995; 95DE-01043628.
XX
XX 24-NOV-1995; 95DE-01043628.
XX
XX (FORS/) FORSSMANN W.
XX
XX Forssmann W, Kist A, Kruhoeffner M, Meyer M, Pardigol A, Heine G;
XX WPI; 1997-290350/27.
XX
XX New guanyl cyclase C activating peptide fragments - have insulinotropic
XX activity, useful for treating diabetes, etc.
XX
XX Claim 3; Fig 3; 33pp; German.
XX
XX Peptides AAW18478-W18483 represent fragments of the guanyl cyclase C
XX activating peptide, GCAP-II, obtained by digestion with trypsin. GCAP-II
XX is involved in insulin secretion by pancreatic beta cells. This peptide
XX fragment could be used to which affects insulin secretion by the beta
XX cells treat pancreatic endocrine disorders, especially diabetes mellitus
XX type II, renal and intestinal disorders, disorders of the respiratory,
XX gastrointestinal and urogenital apparatus, disorders of the
XX cardiovascular and nervous systems, disorders of the integuments and
XX sense organs and diseases associated with GCAP II (89-112) deficiency.
XX This peptide can be used for treatment of electrolyte effects on bone
XX reconstruction (osteoporosis) or the dental apparatus. Antibodies to GCAP
XX -II (89-112) can be used to treat diseases associated with overproduction
XX of GCAP-II (89-112). Human GCAP-II (89-112) and GCAP I (99-15) cDNA are
XX useful for diagnosis and treatment of the above disorders e.g. gene
XX therapy for diabetes
XX
XX Sequence 77 AA;
XX
XX Query Match 70.3%; Score 407; DB 2; Length 77;
XX Best Local Similarity 100.0%; Pred. No. 6.7e-38;
XX Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 VQESMKKLSLEAQWAPSPRLQASLLPAVCHHPALPDQLQPVCAQSEASSIFKTLRTI 95
DB 1 VQESMKKLSLEAQWAPSPRLQASLLPAVCHHPALPDQLQPVCAQSEASSIFKTLRTI 60

Query Match 70.3%; Score 407; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 6.7e-38;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 ANDDCELVCNVACTGCL 112
DB 61 ANDDCELVCNVACTGCL 77

Search completed: March 26, 2005, 16:44:28
Job time : 122.628 secs
```



***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2005, 16:38:19 ; Search time 33.5182 Seconds  
(without alignments)  
249.437 Million cell updates/sec

Title: US-10-775-481A-56  
Perfect score: 579  
Sequence: 1 MGCRAAGSLPGVAVVLL.....RTIANDCELGVNACTGCL 112

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description          |
|------------|-------|-------------|--------|----|----------------------|
| 1          | 162.5 | 28.1        | 115    | 2  | US-07-903-029-4      |
| 2          | 162.5 | 28.1        | 115    | 2  | US-07-903-029-5      |
| 3          | 133   | 23.0        | 115    | 2  | US-07-903-029-6      |
| 4          | 96    | 16.6        | 16     | 1  | US-08-145-940-1      |
| 5          | 96    | 16.6        | 16     | 2  | US-08-583-447A-56    |
| 6          | 90    | 15.5        | 15     | 1  | US-08-145-940-2      |
| 7          | 81    | 14.0        | 15     | 2  | US-08-583-447A-55    |
| 8          | 80.5  | 13.9        | 1206   | 4  | US-09-252-991A-19632 |
| 9          | 75    | 13.0        | 467    | 4  | US-09-522-775A-4     |
| 10         | 74.5  | 12.9        | 623    | 4  | US-09-949-016-7573   |
| 11         | 74.5  | 12.9        | 800    | 2  | US-08-469-537A-72    |
| 12         | 74.5  | 12.9        | 800    | 2  | US-08-469-537A-78    |
| 13         | 74.5  | 12.9        | 937    | 2  | US-08-469-537A-105   |
| 14         | 73.5  | 12.7        | 208    | 4  | US-09-328-352-4746   |
| 15         | 72    | 12.4        | 467    | 4  | US-09-522-775A-2     |
| 16         | 72    | 12.4        | 475    | 4  | US-09-949-016-10456  |
| 17         | 71.5  | 12.3        | 301    | 3  | US-08-829-525-24     |
| 18         | 71.5  | 12.3        | 301    | 3  | US-08-609-583A-24    |
| 19         | 71.5  | 12.3        | 301    | 3  | US-08-937-399-24     |
| 20         | 71.5  | 12.3        | 301    | 4  | US-09-310-367-24     |
| 21         | 71.5  | 12.3        | 301    | 4  | US-09-032-337-24     |
| 22         | 71.5  | 12.3        | 301    | 4  | US-09-464-231-24     |
| 23         | 71    | 12.3        | 88     | 4  | US-09-252-991A-24612 |
| 24         | 71    | 12.3        | 260    | 4  | US-09-252-991A-29887 |
| 25         | 69.5  | 12.0        | 309    | 4  | US-09-489-039A-12224 |
| 26         | 69.5  | 12.0        | 350    | 4  | US-09-270-767-58582  |
| 27         | 69.5  | 12.0        | 559    | 4  | US-09-270-767-43239  |

|    |      |      |      |   |                      |                   |
|----|------|------|------|---|----------------------|-------------------|
| 28 | 69.5 | 12.0 | 934  | 4 | US-09-949-016-7286   | Sequence 7286, Ap |
| 29 | 68.5 | 11.8 | 1046 | 4 | US-09-489-039A-8148  | Sequence 8148, Ap |
| 30 | 68   | 11.7 | 242  | 4 | US-09-059-625-74     | Sequence 74, Appl |
| 31 | 68   | 11.7 | 376  | 4 | US-09-902-540-15516  | Sequence 15516, A |
| 32 | 67.5 | 11.7 | 680  | 4 | US-09-949-016-10770  | Sequence 10770, A |
| 33 | 67.5 | 11.7 | 953  | 4 | US-09-902-540-14681  | Sequence 14681, A |
| 34 | 67   | 11.6 | 475  | 4 | US-09-252-991A-21041 | Sequence 21041, A |
| 35 | 66   | 11.4 | 376  | 4 | US-09-270-767-42614  | Sequence 42614, A |
| 36 | 64.5 | 11.1 | 332  | 2 | US-08-958-642-2      | Sequence 2, Appli |
| 37 | 64.5 | 11.1 | 332  | 3 | US-08-778-394-4      | Sequence 4, Appli |
| 38 | 64.5 | 11.1 | 332  | 3 | US-08-778-423A-2     | Sequence 2, Appli |
| 39 | 64.5 | 11.1 | 488  | 1 | US-07-794-393-2      | Sequence 2, Appli |
| 40 | 64.5 | 11.1 | 488  | 1 | US-08-001-711-2      | Sequence 2, Appli |
| 41 | 64.5 | 11.1 | 488  | 3 | US-08-704-711A-22    | Sequence 22, Appl |
| 42 | 64.5 | 11.1 | 488  | 3 | US-09-521-220-22     | Sequence 22, Appl |
| 43 | 64.5 | 11.1 | 488  | 3 | US-09-391-104-31     | Sequence 31, Appl |
| 44 | 64.5 | 11.1 | 489  | 3 | US-08-448-489-11     | Sequence 11, Appl |
| 45 | 64.5 | 11.1 | 489  | 4 | US-09-689-730-11     | Sequence 11, Appl |

ALIGNMENTS

RESULT 1  
US-07-903-029-4  
; Sequence 4, Application US/07903029  
; Patent No. 5969097  
; GENERAL INFORMATION:  
; APPLICANT: Wiegand, Roger C.  
; APPLICANT: Currie, Mark C.  
; TITLE OF INVENTION: Human Guanylin  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dennis A. Bennett, Monsanto Co., A3SG  
; STREET: 800 N. Lindbergh Blvd.  
; CITY: St. Louis  
; STATE: Missouri  
; COUNTRY: USA  
; ZIP: 63167  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/903,029  
; FILING DATE: 19920623  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bennett, Dennis A.  
; REGISTRATION NUMBER: 34,547  
; REFERENCE/DOCKET NUMBER: 07-21(872)A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314)694-5402  
; TELEFAX: (314)694-9009  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 115 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-903-029-4

|                       |        |                                 |          |                           |        |        |     |      |    |
|-----------------------|--------|---------------------------------|----------|---------------------------|--------|--------|-----|------|----|
| Query Match           | 28.1%; | Score                           | 162.5;   | DB 2;                     | Length | 115;   |     |      |    |
| Best Local Similarity | 36.1%; | Pred. No.                       | 1.1e-12; |                           |        |        |     |      |    |
| Matches               | 35;    | Conservative                    | 15;      | Mismatches                | 32;    | Indels | 15; | Gaps | 3; |
| Qy                    | 27     | VTIQYQCFRVLQSMKKLSDL            | -----    | AQWAPSRLQAGSLLPVCHHPALPDL | 76     |        |     |      |    |
| Db                    | 22     | VTQDGNFSLSLSVKLKKDLPQPRVKLRNFAP | ---      | GEPPVFLCSNPFPEEL          | 78     |        |     |      |    |
| Qy                    | 77     | QPVCAQSEASSIFKTLRTIAND          | --       | DCELCVNACTGC              | 111    |        |     |      |    |



TELEFAX: (708)470-6881  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-145-940-1

Query Match 16.6%; Score 96; DB 1; Length 16;  
Best Local Similarity 100.0%; Pred. No. 1.8e-05;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 97 NDCCLCVNVACTGCL 112  
Db 1 NDCCLCVNVACTGCL 16

## RESULT 5

US-08-583-447A-56  
; Sequence 56, Application US/08583447A  
; Patent No. 5879656  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: ST Receptor Binding Compounds and  
; TITLE OF INVENTION: Methods of Using the Same  
; NUMBER OF SEQUENCES: 56  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5879656ris  
; STREET: One Liberty Place, 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: WordPerfect 6.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/583,447A  
; FILING DATE: 05-JAN-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/141,892  
; FILING DATE: 26-OCT-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: TJU-1702  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 56:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-583-447A-56

Query Match 16.6%; Score 96; DB 2; Length 16;  
Best Local Similarity 100.0%; Pred. No. 1.8e-05;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 97 NDCCLCVNVACTGCL 112  
Db 1 NDCCLCVNVACTGCL 16

## RESULT 6

US-08-145-940-2

; Sequence 2, Application US/08145940  
; Patent No. 5489670  
; GENERAL INFORMATION:  
; APPLICANT: Currie, Mark G.  
; APPLICANT: Kita, Toshihiro  
; APPLICANT: Smith, Christine E.  
; APPLICANT: Fok, Kam F.  
; TITLE OF INVENTION: Human Uroguanylin  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,  
; ADDRESSEE: Corporate Patent Dept.  
; STREET: P. O. Box 5110  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60680  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/145,940  
; FILING DATE:  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bennett, Dennis A.  
; REGISTRATION NUMBER: 34,547  
; REFERENCE/DOCKET NUMBER: 07-21(808)A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (708)470-6501  
; TELEFAX: (708)470-6881  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-145-940-2

Query Match 15.5%; Score 90; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 9.5e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 98 DDCELNVNVACTGCL 112  
Db 1 DDCELNVNVACTGCL 15

## RESULT 7

US-08-583-447A-55  
; Sequence 55, Application US/08583447A  
; Patent No. 5879656  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: ST Receptor Binding Compounds and  
; TITLE OF INVENTION: Methods of Using the Same  
; NUMBER OF SEQUENCES: 56  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5879656ris  
; STREET: One Liberty Place, 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: WordPerfect 6.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/583,447A

; FILING DATE: 05-JAN-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/141,892  
; FILING DATE: 26-OCT-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DeLuca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: TTU-1702  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 55:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-583-447A-55

Query Match 14.0%; Score 81; DB 2; Length 15;  
Best Local Similarity 85.7%; Pred. No. 0.0013;  
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 98 DCELCVNACTGC 111  
DB 2 EDCELCINACTGC 15

## RESULT 8

US-09-252-991A-19632  
; Sequence 19632, Application US/09252991A  
; Patent No. 6551795

; GENERAL INFORMATION:  
; APPLICANT: MARC J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 19632  
; LENGTH: 1206  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-19632

Query Match 13.9%; Score 80.5; DB 4; Length 1206;  
Best Local Similarity 36.5%; Pred. No. 0.68;  
Matches 27; Conservative 6; Mismatches 34; Indels 7; Gaps 3;

QY 33 GFRVQLESKKUSDLERQAWSPRLQAQSLLPVACHHPALPDLPVCASQEAASSIFKTL 92  
DB 853 GFSLOAEGGFKLADPQLAWRP---LQGWRLPTVLHAP---DLFG-CGHYRVIQPFSTL 905  
QY 93 RTIANDDCELCVNV 106  
DB 906 RESASIDGALSTGL 919

## RESULT 9

US-09-522-775A-4  
; Sequence 4, Application US/09522775A  
; Patent No. 6660906  
; GENERAL INFORMATION:  
; APPLICANT: Tschlis, Philip N.  
; TITLE OF INVENTION: Inhibition of Tpl2 To Treat Inflammatory Diseases  
; FILE REFERENCE: OTT-3202

; CURRENT APPLICATION NUMBER: US/09/522,775A  
; CURRENT FILING DATE: 2000-03-08  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 4  
; LENGTH: 467  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Rat Tpl2 polypeptide sequence  
US-09-522-775A-4

Query Match 13.0%; Score 75; DB 4; Length 467;  
Best Local Similarity 33.3%; Pred. No. 0.9;  
Matches 25; Conservative 15; Mismatches 25; Indels 10; Gaps 4;

QY 41 MKKLSLEAQAQWAPSPRLQAQSLLPVACHHPAL--PDLPVCASQEAASSIFKTLRTIAND 98  
DB 361 MRELIEAALERNPNHRPKAADLL---KHEALNPREDQPCQSLD-SALFDRKLLSRK 415  
QY 99 DCELCVNVNVA---CTG 110  
DB 416 ELELPENIADSSCTG 430

## RESULT 10

US-09-949-016-7573  
; Sequence 7573, Application US/09949016  
; Patent No. 6812339

; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7573  
; LENGTH: 623  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-7573

Query Match 12.9%; Score 74.5; DB 4; Length 623;  
Best Local Similarity 24.1%; Pred. No. 1.6;  
Matches 26; Conservative 20; Mismatches 29; Indels 33; Gaps 6;

QY 12 GVAVVLLLLQSTQSVYIQGFRVQLESF-----KKLSLEAQAQWAPSPRLQAQ 60  
DB 144 GIACARFI---GNRTVWESLHMQGEIENQITAAFTWGTSSHLSDKCSQA----- 192  
QY 61 SLLPAVCHHPALPDLPVCASQEAASSIFKTLRTIANDDCELCVNVAC 108  
DB 193 --IFSLCHYAF-----PYC--DETSSVPKP-RDLCRDECEILENVLC 229

## RESULT 11

US-08-469-537A-72  
; Sequence 72, Application US/08469537A  
; Patent No. 5843749  
; GENERAL INFORMATION:  
; APPLICANT: Maisondier, et al.  
; TITLE OF INVENTION: EHK AND ROR TYROSINE  
; TITLE OF INVENTION: KINASES  
; NUMBER OF SEQUENCES: 107  
; CORRESPONDENCE ADDRESS:





```
; NAME: Kempler, Ph.D., Gail M
; REGISTRATION NUMBER: 32,143
; REFERENCE/DOCKET NUMBER: REG 070C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 914-345-7400
; TELEFAX: 914-345-7721
; TELEX:
; INFORMATION FOR SEQ ID NO: 105:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 937 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Human ROR1
; LOCATION: 1...937
; OTHER INFORMATION:
; US-08-469-537A-105

Query Match      12.9%; Score 74.5; DB 2; Length 937;
Best Local Similarity 24.1%; Pred. No. 2.7;
Matches 26; Conservative 20; Mismatches 29; Indels 33; Gaps 6;

QY 12 GVAVVLLLLQSTQSVYIQGFRVQLES-----KKLSDLQAOWAPSPRLQAO 60
Db 175 GIACARI--GNETVMSLHWGEIENQITAAFTWIGTSSHLSDKCSQFA----- 223

QY 61 SLPLAVCHHPALPODQPCVCAQSEASSIFKTLRTIANDDCELCVNVAC 108
Db 224 --IPSLCHYAF-----PYC--DETSSVPKP-RDLCRDECEILENVLC 260

RESULT 14
US-09-328-352-4746
; Sequence 4746, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4746
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-4746

Query Match      12.7%; Score 73.5; DB 4; Length 208;
Best Local Similarity 31.7%; Pred. No. 0.45;
Matches 26; Conservative 12; Mismatches 35; Indels 9; Gaps 4;

QY 18 LLLQSTQSVYI--YQGFRVQLESMMKLSDLQAQ--WAPSPRLQAOSSLPAVC---HHP 70
Db 70 LALLQGRSLIELIETAGGFRFQVRS--KYRNIIAQIWPDPRTKLSPSLLETAVIAYHQP 127

QY 71 ALPDQLQPCVCAQSEASSIFKTL 92
Db 128 VTRADIEQIRGVSNNSQILRTL 149

RESULT 15
US-09-522-775A-2
; Sequence 2, Application US/09522775A
; Patent No. 6660906
; GENERAL INFORMATION:
; APPLICANT: Tsichlis, Philip N.
; TITLE OF INVENTION: Inhibition of Tpl2 To Treat Inflammatory Diseases
; FILE REFERENCE: OTT-3202
```

```
; CURRENT APPLICATION NUMBER: US/09/522,775A
; CURRENT FILING DATE: 2000-03-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human Tpl2 polypeptide sequence
; US-09-522-775A-2
```

```
Query Match      12.4%; Score 72; DB 4; Length 467;
Best Local Similarity 32.0%; Pred. No. 2.2;
Matches 24; Conservative 16; Mismatches 25; Indels 10; Gaps 4;

QY 41 MKKLSDLQAOWAPSPRLQAOSSLPAVCHHPAL--PDQLQPCVCAQSEASSIFKTLRTIAN 98
Db 361 MRELIEASLERNPVHRPRAADLL-----KHEALNPREDQPCQSLD--SALLERKRLLSRK 415

QY 99 DCELCVNVA---CTG 110
Db 416 ELELPENIADSSCTG 430
```

Search completed: March 26, 2005, 16:51:13  
Job time : 34.5182 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2005, 16:48:40 ; Search time 111.591 Seconds  
(without alignments)  
332.314 Million cell updates/sec

Title: US-10-775-481A-56  
Perfect score: 579  
Sequence: 1 MGCRAASGLPGVAVVLLLL.....RTIANDCCLCVNVACTGCL 112

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1407402 seqs, 331100923 residues

Total number of hits satisfying chosen parameters: 1407402

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 579   | 100.0       | 112    | 17    | US-10-479-606-5    |
| 2          | 252   | 43.5        | 109    | 17    | US-10-479-606-6    |
| 3          | 216   | 37.3        | 85     | 9     | US-09-925-299-1527 |
| 4          | 216   | 37.3        | 85     | 10    | US-09-925-299-1527 |
| 5          | 162.5 | 28.1        | 115    | 9     | US-09-981-353-61   |
| 6          | 162.5 | 28.1        | 115    | 14    | US-10-235-994-22   |
| 7          | 162.5 | 28.1        | 115    | 14    | US-10-262-473-12   |
| 8          | 162.5 | 28.1        | 115    | 17    | US-10-479-606-4    |
| 9          | 136   | 23.5        | 108    | 14    | US-10-262-473-16   |
| 10         | 132   | 22.8        | 102    | 14    | US-10-262-473-14   |
| 11         | 96    | 16.6        | 16     | 14    | US-10-107-814-1    |
| 12         | 96    | 16.6        | 16     | 14    | US-10-197-954-141  |
| 13         | 96    | 16.6        | 16     | 15    | US-10-621-684-56   |

|    |      |      |      |    |                    |                   |
|----|------|------|------|----|--------------------|-------------------|
| 14 | 96   | 16.6 | 16   | 17 | US-10-479-606-2    | Sequence 2, Appli |
| 15 | 96   | 16.6 | 16   | 17 | US-10-760-085-141  | Sequence 141, App |
| 16 | 92   | 15.9 | 16   | 14 | US-10-107-814-20   | Sequence 20, Appl |
| 17 | 91   | 15.7 | 87   | 9  | US-09-925-299-1162 | Sequence 1162, Ap |
| 18 | 91   | 15.7 | 87   | 10 | US-09-925-299-1162 | Sequence 1162, Ap |
| 19 | 81   | 14.0 | 15   | 15 | US-10-621-684-55   | Sequence 55, Appl |
| 20 | 80   | 13.8 | 14   | 14 | US-10-107-814-21   | Sequence 21, Appl |
| 21 | 77   | 13.3 | 1010 | 14 | US-10-238-075-945  | Sequence 945, App |
| 22 | 75   | 13.0 | 467  | 9  | US-09-374-579-2    | Sequence 2, Appli |
| 23 | 75   | 13.0 | 467  | 15 | US-10-374-414-2    | Sequence 2, Appli |
| 24 | 74.5 | 12.9 | 937  | 9  | US-09-974-298-129  | Sequence 129, App |
| 25 | 74.5 | 12.9 | 937  | 17 | US-10-482-029-245  | Sequence 245, App |
| 26 | 74   | 12.8 | 72   | 17 | US-10-796-719-21   | Sequence 21, Appl |
| 27 | 73   | 12.6 | 467  | 9  | US-09-374-579-4    | Sequence 4, Appli |
| 28 | 73   | 12.6 | 467  | 15 | US-10-374-414-4    | Sequence 4, Appli |
| 29 | 73   | 12.6 | 467  | 15 | US-10-262-511-228  | Sequence 228, App |
| 30 | 73   | 12.6 | 467  | 15 | US-10-287-226-2    | Sequence 2, Appli |
| 31 | 73   | 12.6 | 467  | 17 | US-10-753-267-106  | Sequence 106, App |
| 32 | 72   | 12.4 | 351  | 16 | US-10-664-421-155  | Sequence 155, App |
| 33 | 72   | 12.4 | 431  | 15 | US-10-262-511-230  | Sequence 230, App |
| 34 | 72   | 12.4 | 467  | 15 | US-10-114-270-112  | Sequence 112, App |
| 35 | 72   | 12.4 | 467  | 15 | US-10-262-511-232  | Sequence 232, App |
| 36 | 72   | 12.4 | 467  | 15 | US-10-287-226-24   | Sequence 24, Appl |
| 37 | 72   | 12.4 | 467  | 15 | US-10-287-226-36   | Sequence 36, Appl |
| 38 | 72   | 12.4 | 468  | 15 | US-10-287-226-8    | Sequence 8, Appli |
| 39 | 72   | 12.4 | 471  | 15 | US-10-287-226-18   | Sequence 18, Appl |
| 40 | 72   | 12.4 | 474  | 15 | US-10-287-226-4    | Sequence 4, Appli |
| 41 | 72   | 12.4 | 474  | 15 | US-10-287-226-26   | Sequence 26, Appl |
| 42 | 72   | 12.4 | 475  | 15 | US-10-287-226-16   | Sequence 16, Appl |
| 43 | 72   | 12.4 | 475  | 15 | US-10-287-226-32   | Sequence 32, Appl |
| 44 | 72   | 12.4 | 478  | 15 | US-10-287-226-10   | Sequence 10, Appl |
| 45 | 72   | 12.4 | 486  | 15 | US-10-287-226-20   | Sequence 20, Appl |

ALIGNMENTS

RESULT 1

US-10-479-606-5  
; Sequence 5, Application US/10479606  
; Publication No. US20050032684A1  
; GENERAL INFORMATION:  
; APPLICANT: Savas, Yulcel  
; TITLE OF INVENTION: Guanylate-cyclase C ligand, administered via the airways, for the  
; TITLE OF INVENTION: treatment of respiratory airway problems  
; FILE REFERENCE: 03100192aa  
; CURRENT APPLICATION NUMBER: US/10/479,606  
; PRIOR FILING DATE: 2003-12-04  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: PCT/DE02/02040  
; PRIOR FILING DATE: 2002-06-05  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 5  
; LENGTH: 112  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-479-606-5

Query Match 100.0%; Score 579; DB 17; Length 112;  
Best Local Similarity 100.0%; Pred. No. 8.2e-59;  
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGCRAASGLPGVAVVLLLLQSTQSVYIQGFRVQLESMMKKLSDLLEAOWAPSPRIQAQ 60

Db 1 MGCRAASGLPGVAVVLLLLQSTQSVYIQGFRVQLESMMKKLSDLLEAOWAPSPRIQAQ 60

Qy 61 SLIPAVCHHPALPDQLQPVCAQSEASSIFKTLRTIANDCCLCVNVACTGCL 112

Db 61 SLIPAVCHHPALPDQLQPVCAQSEASSIFKTLRTIANDCCLCVNVACTGCL 112

```
RESULT 2
US-10-479-606-6
; Sequence 6, Application US/10479606
; Publication No. US20050032684A1
; GENERAL INFORMATION:
; APPLICANT: Cetin, Yalcin
; APPLICANT: Savas, Yuksel
; TITLE OF INVENTION: Guanylate-cyclase C ligand, administered via the airways, for the
; FILE REFERENCE: 03100192aa
; CURRENT APPLICATION NUMBER: US/10/479,606
; CURRENT FILING DATE: 2003-12-04
; PRIOR APPLICATION NUMBER: DE10127119.0
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/DE02/02040
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 6
; LENGTH: 109
; TYPE: PRT
; ORGANISM: opossum
US-10-479-606-6

Query Match 43.5%; Score 252; DB 17; Length 109;
Best Local Similarity 50.5%; Pred. No. 4.9e-21;
Matches 52; Conservative 17; Mismatches 32; Indels 2; Gaps 2;

Qy 10 LP-GVAVVLLLLQSTQSVYIQGFRVQLESMMKLSDLBAQWAPSPRLQAQSELLPAV-C 67
Db 6 LPMAVTAMLLILAQNTQSVYIQGFRVQLESMMKLSDLBAQWAPSPRLQAQSELLPAV-C 65

Qy 68 HHPALPDLPQVCASQASISFKTLRTIANDCCLCVNACTG 110
Db 66 SDPALPSDLPVCENSQAVNIFRALRYINQECCLCNMACTG 108

RESULT 3
US-09-925-299-1527
; Sequence 1527, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1527
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (9)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (28)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (29)
```

```
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (31)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (32)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (36)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (39)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (42)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (43)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (47)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (52)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (53)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (55)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (60)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (66)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (70)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (71)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (83)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-1527

Query Match 37.3%; Score 216; DB 9; Length 85;
Best Local Similarity 62.5%; Pred. No. 5.1e-17;
Matches 45; Conservative 3; Mismatches 24; Indels 0; Gaps 0;

Qy 1 MGCRAASGLPGVAVVLLLLQSTQSVYIQGFRVQLESMMKLSDLBAQWAPSPRLQAQ 60
Db 11 MGCWVASGLLTGLAEVLXLLXLTXXQGXQYXXFRVLESMMXLLXDLBAWAPSPXLEAX 70

Qy 61 SILLPAVCHHPAL 72
Db 71 XLLAAVCHHPAL 82

RESULT 4
US-09-925-299-1527
; Sequence 1527, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
```

```

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PAI02
; CURRENT APPLICATION NUMBER: US/09/925,299
; PRIOR FILING DATE: 2001-08-10
; PCT APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PCT APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; LENGTH: 85
; TYPE: PRP
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (9)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (28)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (29)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (31)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (32)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (36)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (39)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (42)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (43)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (47)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (52)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (53)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (55)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (60)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (66)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE

```

---

; LOCATION: (70)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (71)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (83)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-299-1527

Query Match      37.3%;    Score 216; DB 10; Length 85;  
Best Local Similarity    62.5%; Pred.No. 5.1e-17;  
Matches     45; Conservative       3; Mismatches    24; Indels     0; Gaps     0

Qy                  1    MGCRAASGLPGVAVVVLLLLLLOSTQSVTYIQGFRVQLSEMKKLSLEAQWAPSPRLQAQ 60  
                     ||||| | : ||| : ||| : ||| : ||| : ||| : ||| : ||| :  
Db                  11    MCGWVASGLLTGLAEVLXXXTXGXKKXXFKRVKLSEMXLXLDEAXWAPSXPLEAX 70  
                     || ||| | : ||| : ||| : ||| : ||| : ||| : ||| :  
Qy                  61    SLPAVCCHPAL 72  
                     || ||| | : ||| : ||| : ||| : ||| : ||| : ||| :  
Db                  71    XLAAVCHHPAL 82

RESULT 5  
US-09-981-353-61  
; Sequence 61, Application US/09981353  
; Patent No. US20020160382A1  
; GENERAL INFORMATION:  
; APPLICANT: Lasek, Amy W.  
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER  
; FILE REFERENCE: PA-0038 US  
; CURRENT APPLICATION NUMBER: US/09/981,353  
; CURRENT FILING DATE: 2001-10-11  
; NUMBER OF SEQ ID NOS: 194  
; SOFTWARE: PERL Program  
; SEQ ID NO 61  
; LENGTH: 115  
; TYPE: PRP  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. US20020160382A1 1800311CD1  
US-09-981-353-61

Query Match      28.1%;    Score 162.5; DB 9; Length 115;  
Best Local Similarity    36.1%; Pred.No. 1.2e-10;  
Matches     35; Conservative       15; Mismatches    32; Indels     15; Gaps     3

Qy                  27    VTIYQGFRVQLSEMKKLSDSL-----AQWAPSPRLQAQSILLPAVCHHPALPDOL 76  
                     ||||| | : ||| : ||| : ||| : ||| : ||| : ||| : ||| :  
Db                  22    VTVDGNFSLFSVKLDLQEPQPRVGKLRNFAPIP---GEFPVPILCSNPNPEEL 78  
                     ||||| | : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

Qy                  77    QPVCAQSQBASSIFKTIRTIAND--DCELCVNVACTGC 111  
                     : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :  
Db                  79    KPLCKEPPNAQEILQRLEEIAEDPGTCEICAYAACTGC 115  
                     : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

RESULT 6  
US-10-235-994-22  
; Sequence 22, Application US/10235994  
; Publication No. US20030101002A1  
; GENERAL INFORMATION:  
; APPLICANT: Bartha, Gabor  
; APPLICANT: Walker, Michael  
; TITLE OF INVENTION: METHODS FOR ANALYZING GENE EXPRESSION PATTERNS  
; FILE REFERENCE: ICYP012  
; CURRENT APPLICATION NUMBER: US/10/235,994  
; CURRENT FILING DATE: 2002-09-04  
; PRIOR APPLICATION NUMBER: US/10/003,608  
; PRIOR FILING DATE: 2001-11-01  
; PRIOR APPLICATION NUMBER: 60/245,081  
; PRIOR FILING DATE: 2000-11-01

```

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PAI02
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1527
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (9)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (28)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (29)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (31)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (32)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (36)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (39)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (42)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (43)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (47)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (52)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (53)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (55)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (60)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (66)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE

```

```
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Human
US-10-235-994-22
```

```
Query Match      28.1%; Score 162.5; DB 14; Length 115;
Best Local Similarity 36.1%; Pred. No. 1.2e-10;
Matches 35; Conservative 15; Mismatches 32; Indels 15; Gaps 3;

QY 27 VYIQGFRVQLSESMKKLSLE-----AOWAPSPRLQAQSLLPVCHHPALPQDL 76
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 22 VTVQDGNFSFSLSESVKLLKDLQEPQPRVGLKRNFAPIP---GEPVPLCSNPNFPFEL 78
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 77 QPVCASQASSTFKTLRTIAND--DCELCVNVACTGC 111
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 79 KPLCKEPNAQEILQRLLEEIAEDPGTCEICAYAACTGC 115
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

## RESULT 7

US-10-262-473-12

```
; Sequence 12, Application US/10262473
; Publication No. US20030199442A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook, John,
; APPLICANT: Burgess, Catherine,
; APPLICANT: Gorman, Linda,
; APPLICANT: Guo, Xiaojia,
; APPLICANT: Lepley, Denise,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rastelli, Luca,
; APPLICANT: Reiger, Daniel,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-462B
; CURRENT APPLICATION NUMBER: US/10/262,473
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/349,575
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 12
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-473-12
```

```
Query Match      28.1%; Score 162.5; DB 14; Length 115;
Best Local Similarity 36.1%; Pred. No. 1.2e-10;
Matches 35; Conservative 15; Mismatches 32; Indels 15; Gaps 3;

QY 27 VYIQGFRVQLSESMKKLSLE-----AOWAPSPRLQAQSLLPVCHHPALPQDL 76
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 22 VTVQDGNFSFSLSESVKLLKDLQEPQPRVGLKRNFAPIP---GEPVPLCSNPNFPFEL 78
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 77 QPVCASQASSTFKTLRTIAND--DCELCVNVACTGC 111
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 79 KPLCKEPNAQEILQRLLEEIAEDPGTCEICAYAACTGC 115
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

## RESULT 8

US-10-479-606-4

```
; Sequence 4, Application US/10479606
; Publication No. US20050032684A1
; GENERAL INFORMATION:
; APPLICANT: Cetin, Yalcin
; APPLICANT: Savas, Yuksel
; TITLE OF INVENTION: Guanylate-cyclase C ligand, administered via the airways, for the
; FILE REFERENCE: 03100192aa
; CURRENT APPLICATION NUMBER: US/10/479,606
; CURRENT FILING DATE: 2003-12-04
; PRIOR APPLICATION NUMBER: DE10127119.0
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/DE02/02040
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 115
; TYPE: PRT
; ORGANISM: rat or homo sapiens (guanylin)
US-10-479-606-4
```

```
Query Match      28.1%; Score 162.5; DB 17; Length 115;
Best Local Similarity 36.1%; Pred. No. 1.2e-10;
Matches 35; Conservative 15; Mismatches 32; Indels 15; Gaps 3;

QY 27 VYIQGFRVQLSESMKKLSLE-----AOWAPSPRLQAQSLLPVCHHPALPQDL 76
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 22 VTVQDGNFSFSLSESVKLLKDLQEPQPRVGLKRNFAPIP---GEPVPLCSNPNFPFEL 78
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 77 QPVCASQASSTFKTLRTIAND--DCELCVNVACTGC 111
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 79 KPLCKEPNAQEILQRLLEEIAEDPGTCEICAYAACTGC 115
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

## RESULT 9

```
US-10-262-473-16
; Sequence 16, Application US/10262473
; Publication No. US20030199442A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook, John,
; APPLICANT: Burgess, Catherine,
; APPLICANT: Gorman, Linda,
; APPLICANT: Guo, Xiaojia,
; APPLICANT: Lepley, Denise,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rastelli, Luca,
; APPLICANT: Reiger, Daniel,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-462B
; CURRENT APPLICATION NUMBER: US/10/262,473
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/349,575
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 16
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-473-16
```

Query Match 23.5%; Score 136; DB 14; Length 108;

Best Local Similarity 36.0%; Pred. No. 1.2e-07;  
Matches 32; Conservative 15; Mismatches 32; Indels 10; Gaps 4;  
Qy 27 VYIQGFRVQLE-SMKKLSLEAQWAPSRPRLQAQSLLPVACHHPALPDQLQPVCAQSEA 85  
Db 25 VTVDQGNFSFSLPRVGLRN----FAPIP---GEPVVPILCSNPNFPBELKPLCKEPNA 77  
Qy 86 SSIFKTLRTIAND--DCELCNVNACTGCL 112  
Db 78 QEILQRLLEEIAEDPGTCETICAYAACTGC 106  
RESULT 10  
US-10-262-473-14  
; Sequence 14, Application US/10262473  
; Publication No. US20030199442A1  
; GENERAL INFORMATION:  
; APPLICANT: Alsobrook, John,  
; APPLICANT: Burgess, Catherine,  
; APPLICANT: Gorman, Linda,  
; APPLICANT: Guo, Xiaojia,  
; APPLICANT: Lepley, Denise,  
; APPLICANT: Patturajan, Meera,  
; APPLICANT: Rastelli, Luca,  
; APPLICANT: Reiger, Daniel,  
; APPLICANT: Spytek, Kimberly,  
; APPLICANT: Zhong, Mei  
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
; FILE REFERENCE: 21402-462B  
; CURRENT APPLICATION NUMBER: US/10/262,473  
; CURRENT FILING DATE: 2003-01-28  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/327,917  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/328,029  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/328,056  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/349,575  
; PRIOR FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 60/381,038  
; PRIOR FILING DATE: 2002-05-16  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Curaseqlist version 0.1  
; SEQ ID NO 14  
; LENGTH: 102  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-262-473-14  
Query Match 22.8%; Score 132; DB 14; Length 102;  
Best Local Similarity 35.2%; Pred. No. 3.3e-07;  
Matches 31; Conservative 15; Mismatches 32; Indels 10; Gaps 4;  
Qy 27 VYIQGFRVQLE-SMKKLSLEAQWAPSRPRLQAQSLLPVACHHPALPDQLQPVCAQSEA 85  
Db 22 VTVDQGNFSFSLPRVGLRN----FAPIP---GEPVVPILCSNPNFPBELKPLCKEPNA 74  
Qy 86 SSIFKTLRTIAND--DCELCNVNACTGC 111  
Db 75 QEILQRLLEEIAEDPGTCETICAYAACTGC 102  
RESULT 11  
US-10-107-814-1  
; Sequence 1, Application US/10107814  
; Publication No. US20030073628A1  
; GENERAL INFORMATION:  
; APPLICANT: SHAILUBHAI, KUNWAR  
; APPLICANT: NIKIFOROVICH, GREGORY  
; APPLICANT: JACOB, GARY S.  
; TITLE OF INVENTION: GUANYLATE CYCLASE RECEPTOR AGONISTS FOR THE TREATMENT  
; TITLE OF INVENTION: OF TISSUE INFLAMMATION AND CARCINOGENESIS  
; FILE REFERENCE: 81361/284943/MAS

CURRENT APPLICATION NUMBER: US/10/107,814  
CURRENT FILING DATE: 2002-03-28  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Homo sapiens  
NAME/KEY: DISULFID  
LOCATION: (4)..(12)  
NAME/KEY: DISULFID  
LOCATION: (7)..(15)  
US-10-107-814-1  
Query Match 16.6%; Score 96; DB 14; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.00043;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 97 NDDCELCNVNACTGCL 112  
Db 1 NDDCELCNVNACTGCL 16  
RESULT 12  
US-10-197-954-141  
; Sequence 141, Application US/10197954  
; Publication No. US20030119021A1  
; GENERAL INFORMATION:  
; APPLICANT: K'ater, Hubert  
; APPLICANT: Siddiqi, Suhail  
; APPLICANT: Little, Daniel  
; TITLE OF INVENTION: Capture Compounds, Collections Thereof  
; TITLE OF INVENTION: And Methods For Analyzing The Proteome And Complex  
; TITLE OF INVENTION: Compositions  
; FILE REFERENCE: 24743-2305  
; CURRENT APPLICATION NUMBER: US/10/197,954  
; CURRENT FILING DATE: 2002-07-16  
; PRIOR APPLICATION NUMBER: 60/306,019  
; PRIOR FILING DATE: 2001-07-16  
; PRIOR APPLICATION NUMBER: 60/314,123  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: 60/363,433  
; PRIOR FILING DATE: 2002-03-11  
; NUMBER OF SEQ ID NOS: 149  
; SOFTWARE: Fast-Seq For Windows Version 4.0  
; SEQ ID NO 141  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-197-954-141  
Query Match 16.6%; Score 96; DB 14; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.00043;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 97 NDDCELCNVNACTGCL 112  
Db 1 NDDCELCNVNACTGCL 16  
RESULT 13  
US-10-621-684-56  
; Sequence 56, Application US/10621684  
; Publication No. US20040029182A1  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: ST Receptor Binding Compounds and  
; TITLE OF INVENTION: Methods of Using the Same  
; NUMBER OF SEQUENCES: 56  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. US20040029182A1ris  
; STREET: One Liberty Place, 46th Floor

CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: WordPerfect 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/621,694  
FILING DATE: 17-Jul-2003  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/583,447A  
FILING DATE: 05-JAN-1996  
APPLICATION NUMBER: US 08/141,892  
FILING DATE: 26-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Deluca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-1702  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 56:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 56:  
US-10-621-684-56

Query Match 16.6%; Score 96; DB 15; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.00043;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 NDDCELCVNVACTGCL 112  
|||||  
DB 1 NDDCELCVNVACTGCL 16

## RESULT 14

US-10-479-606-2  
Sequence 2, Application US/10479606  
Publication No. US20050032684A1  
GENERAL INFORMATION:  
APPLICANT: Cetin, Yalcin  
APPLICANT: Savas, Yuksel  
TITLE OF INVENTION: Guanylate-cyclase C ligand, administered via the airways, for the treatment of respiratory airway problems  
FILE REFERENCE: 03100192aa  
CURRENT APPLICATION NUMBER: US/10/479,606  
CURRENT FILING DATE: 2003-12-04  
PRIOR APPLICATION NUMBER: DE10127119.0  
PRIOR FILING DATE: 2001-06-05  
PRIOR APPLICATION NUMBER: PCT/DE02/02040  
PRIOR FILING DATE: 2002-06-05  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 2  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-479-606-2

Query Match 16.6%; Score 96; DB 17; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.00043;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 NDDCELCVNVACTGCL 112  
|||||

Db 1 NDDCELCVNVACTGCL 16

## RESULT 15

US-10-760-085-141  
Sequence 141, Application US/10760085  
Publication No. US20050042771A1  
GENERAL INFORMATION:  
APPLICANT: Hubert K'water  
APPLICANT: Daniel Paul Little  
APPLICANT: Suhaib Mahmood Siddiqi  
APPLICANT: Matthew Peter Grealish  
APPLICANT: Subramaniam Marappan  
APPLICANT: Chester Frederick Hassman III  
APPLICANT: Ping Yip  
TITLE OF INVENTION: Capture Compounds, Collections Thereof  
TITLE OF INVENTION: And Methods For Analyzing The Proteome And Complex Compositions  
FILE REFERENCE: 24743-2309  
CURRENT APPLICATION NUMBER: US/10/760,085  
CURRENT FILING DATE: 2004-01-16  
PRIOR APPLICATION NUMBER: 60/441,398  
PRIOR FILING DATE: 2003-01-16  
NUMBER OF SEQ ID NOS: 149  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 141  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-760-085-141

Query Match 16.6%; Score 96; DB 17; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.00043;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 NDDCELCVNVACTGCL 112  
|||||  
DB 1 NDDCELCVNVACTGCL 16

Search completed: March 26, 2005, 17:16:38  
Job time : 111.591 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2005, 16:37:49 ; Search time 23.708 Seconds  
(without alignments)  
454.541 Million cell updates/sec

Title: US-10-775-481A-56  
Perfect score: 579  
Sequence: 1 MGCRAASGLPGVAVVLLLL.....RTIANDDCELCVNVACTGCL 112

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 579   | 100.0       | 112    | 2     | JC4651      |
| 2          | 162.5 | 28.1        | 115    | 1     | A46279      |
| 3          | 161   | 27.8        | 116    | 2     | JC7620      |
| 4          | 153.5 | 26.5        | 116    | 1     | B46279      |
| 5          | 133   | 23.0        | 115    | 1     | JN0318      |
| 6          | 81    | 14.0        | 72     | 1     | QHEC1B      |
| 7          | 76.5  | 13.2        | 500    | 2     | B86312      |
| 8          | 75    | 13.0        | 467    | 2     | A47388      |
| 9          | 74.5  | 12.9        | 937    | 2     | A45082      |
| 10         | 74    | 12.8        | 72     | 1     | QHEC4       |
| 11         | 74    | 12.8        | 467    | 2     | I49609      |
| 12         | 74    | 12.8        | 824    | 2     | I50618      |
| 13         | 72    | 12.4        | 467    | 2     | A48713      |
| 14         | 71    | 12.3        | 737    | 2     | S63453      |
| 15         | 68.5  | 11.8        | 369    | 2     | AG0386      |
| 16         | 68.5  | 11.8        | 699    | 2     | T18984      |
| 17         | 68    | 11.7        | 873    | 1     | TVFVFS      |
| 18         | 66.5  | 11.5        | 538    | 2     | AC1850      |
| 19         | 66.5  | 11.5        | 1127   | 2     | T21635      |
| 20         | 66.5  | 11.5        | 1165   | 2     | T21636      |
| 21         | 66    | 11.4        | 580    | 2     | T43485      |
| 22         | 65.5  | 11.3        | 144    | 2     | C70937      |
| 23         | 65.5  | 11.3        | 322    | 2     | D86760      |
| 24         | 65.5  | 11.3        | 533    | 1     | TVFVFP      |
| 25         | 65    | 11.2        | 710    | 2     | T52573      |
| 26         | 64.5  | 11.1        | 211    | 2     | D82109      |
| 27         | 64.5  | 11.1        | 367    | 2     | F82098      |
| 28         | 64.5  | 11.1        | 378    | 2     | B89588      |
| 29         | 64.5  | 11.1        | 488    | 2     | S13423      |

|    |      |      |      |   |        |
|----|------|------|------|---|--------|
| 30 | 64.5 | 11.1 | 1736 | 2 | T00391 |
| 31 | 64   | 11.1 | 344  | 2 | T12307 |
| 32 | 64   | 11.1 | 456  | 2 | A31857 |
| 33 | 64   | 11.1 | 465  | 2 | S41644 |
| 34 | 64   | 11.1 | 537  | 2 | T28683 |
| 35 | 64   | 11.1 | 614  | 1 | I48385 |
| 36 | 64   | 11.1 | 1040 | 2 | AD0771 |
| 37 | 63.5 | 11.0 | 349  | 2 | JC5827 |
| 38 | 63.5 | 11.0 | 445  | 2 | T50005 |
| 39 | 63.5 | 11.0 | 589  | 2 | T29895 |
| 40 | 63.5 | 11.0 | 707  | 2 | T40070 |
| 41 | 63.5 | 11.0 | 848  | 2 | S24981 |
| 42 | 63.5 | 11.0 | 1088 | 2 | H96747 |
| 43 | 63   | 10.9 | 96   | 2 | A27166 |
| 44 | 63   | 10.9 | 692  | 2 | E90952 |
| 45 | 63   | 10.9 | 692  | 2 | A85801 |

ALIGNMENTS

RESULT 1

JC4651

uroguanylin precursor - human

N:Alternate names: guanylyl cyclase activating peptide II

C:Species: Homo sapiens (man)

C>Date: 10-May-1996 #sequence revision 19-Jul-1996 #text\_change 09-Jul-2004

C:Accession: JC4651; S63702; S68052

R:Miyazato, M.; Nakazato, M.; Yamaguchi, H.; Date, Y.; Kojima, M.; Kangawa, K.; Matsuo, Biochem. Biophys. Res. Commun. 219, 644-648, 1996

A:Title: Cloning and characterization of a cDNA encoding a precursor for human uroguanylin

A:Reference number: JC4651; MUID:96193705; PMID:8605041

A:Accession: JC4651

A:Molecule type: mRNA

A:Residues: 1-112 <MY>

A:Cross-references: UNIPROT:Q16661; GB:U34279; NID:gl336798; PIDN:AA050416.1; PID:gl3367

R:Hill, O.; Cetin, Y.; Cieslak, A.; Maegert, H.J.; Forssmann, W.G.

Biochim. Biophys. Acta 1253, 146-149, 1995

A:Title: A new human guanylate cyclase-activating peptide (GCAP-II, uroguanylin): precursor

A:Reference number: S63702; MUID:96106424; PMID:8519795

A:Accession: S63702

A:Molecule type: mRNA

A:Residues: 1-112 <HIL>

A:Cross-references: EMBL:Z50753; NID:g974823; PIDN:CAA906239.1; PID:g974824

A:Experimental source: tissue colon

R:Hess, R.; Kuhn, M.; Schulz-Knappe, P.; Raida, M.; Fuchs, M.; Klodt, J.; Adermann, K.; FEBS Lett. 374, 34-38, 1995

A:Title: GCAP-II: isolation and characterization of the circulating form of human urogu

A:Reference number: S68052; MUID:96049550; PMID:7589507

A:Accession: S68052

A:Molecule type: protein

A:Residues: 89-99,'X',101-102,'X',104-107,'X',109-110,'X',112 <HES>

C:Comment: This protein, a member of the guanylin peptide family, is an endogenous activ

C:Superfamily: Guanylin

C:Keywords: intestine

F:1-26/Domain: signal sequence #status predicted <SIG>

F:27-112/Product: uroguanylin #status predicted <WAT>

Query Match 100.0%; Score 579; DB 2; Length 112;  
Best Local Similarity 100.0%; Pred. No. 4.1e-54;  
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGCRAASGLPGVAVVLLLLLQSTQSVYIQGFRVQLESMMKSLDLEAOWAPSPRLQAA 60

Db 1 MGCRAASGLPGVAVVLLLLLQSTQSVYIQGFRVQLESMMKSLDLEAOWAPSPRLQAA 60

Qy 61 SLPAVCHHPALPDQLQPVCAQSEASSIFKTLRTIANDDCELCVNVACTGCL 112

Db 61 SLPAVCHHPALPDQLQPVCAQSEASSIFKTLRTIANDDCELCVNVACTGCL 112

RESULT 2

A46279  
guanylin precursor [validated] - human  
C:Species: Homo sapiens (man)  
C>Date: 21-Sep-1993 #sequence\_revision 26-May-1995 #text\_change 09-Jul-2004  
C:Accession: A46279; S29228; S29807  
R:de Sauvage, F.J.; Keshav, S.; Kuang, W.J.; Gillett, N.; Henzel, W.; Goeddel, D.V.  
Proc. Natl. Acad. Sci. U.S.A. 89, 9089-9093, 1992  
A:Title: Precursor structure, expression, and tissue distribution of human guanylin.  
A:Reference number: A46279; MUID:93028409; PMID:1409606  
A:Accession: A46279  
A:Molecule type: mRNA  
A:Residues: 1-115 <DEL>  
A:Cross-references: UNIPROT:Q02747; GB:M951174; NID:G306823; PIDN:AAA58625.1; PID:G306824  
R:Wiegand, R.C.; Kato, J.; Huang, M.D.; Fok, K.F.; Kachur, J.F.; Currie, M.G.  
FEBS Lett. 311, 150-154, 1992  
A:Title: The circulating bioactive form of human guanylin is a high molecular weight peptide.  
A:Reference number: S29807; MUID:93178628; PMID:8095028  
A:Accession: S29807  
A:Molecule type: protein  
A:Residues: 22-68 <KUH>  
A:Experimental source: plasma  
A:Note: amino-terminal sequencing of mature form and molecular weight of mature form by n of the same.  
C:Comment: Guanylin is an endogenous ligand for an intestine-specific receptor guanylyl n of the same receptor.  
C:Genetics:  
A:Gene: GDB:GUCA2  
A:Cross-references: GDB:136460; OMIM:139392  
A:Map position: lp35-1p34  
C:Superfamily: guanylin  
C:Keywords: hormone; intestine  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-115/Product: guanylin #status experimental <MAT>

Query Match 28.1%; Score 162.5; DB 1; Length 115;  
Best Local Similarity 36.1%; Pred. No. 5.3e-10;  
Matches 35; Conservative 15; Mismatches 32; Indels 15; Gaps 3;

Qy 27 VYIQYQGFVQLSMKSLDLE-----AQWAPSPRLQASLLPAVCHHPALPQDL 76  
Db 22 VTVDGNFSLSLSVKKLDQEPQPRVGLKRNFAIP---GEPVVPILCSNPNPPEL 78

Qy 77 QPVCAQEAASSIFKTLRTIAND--DCELCVNVACTGC 111  
Db 79 KPLCKEPNAQEIQLRLEIAEDPGTCEICAYACTGC 115

RESULT 3  
JC7620  
guanylin precursor, long form - European eel  
C:Species: Anguilla anguilla (European eel)  
C>Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 03-Aug-2001  
C:Accession: JC7620  
R:Comrie, M.M.; Cutler, C.P.; Cramb, G.  
Biochem. Biophys. Res. Commun. 281, 1078-1085, 2001  
A:Title: Cloning and expression of guanylin from the European eel (Anguilla anguilla).  
A:Reference number: JC7620; MUID:21139737; PMID:11243845  
A:Accession: JC7620  
A:Molecule type: mRNA  
A:Residues: 1-116 <COM>  
A:Cross-references: GB:AJ301673  
C:Comment: This protein, a member of a family of heat-stable peptides, is a potent extra axis. This peptide signalling system plays a role in osmoregulation in euryhaline teleosts.  
C:Superfamily: guanylin  
C:Keywords: heat-stable protein; osmoregulation

F:1-28/Domain: signal sequence #status predicted <SIG>  
F:29-116/Product: guanylin precursor, long form #status predicted <MAT>  
F:33-39/Region: homologous #status predicted  
F:69-114/Region: highly conserved #status predicted

Query Match 27.8%; Score 161; DB 2; Length 116;  
Best Local Similarity 39.4%; Pred. No. 7.7e-10;  
Matches 43; Conservative 13; Mismatches 39; Indels 14; Gaps 4;

Qy 14 AVVLLLL-----LQSTQSVYIQGFVQLSMKSLDLEAQWAP-----PRLQASLLPA 65  
Db 12 AVVLLLLACFCFLQESGVWMDGLDLPFLAVKVLKHLGANTMTSPHPNIGSH-----A 67

Qy 66 VCHHPALPDLPQVCAQEAASSIFKTLRTIAN--DCELCVNVACTGC 112  
Db 68 VCSNPLHPAEFLPVCEGASALFNRLVDIITPPDPCEICANAACTGC 116

RESULT 4  
B46279  
guanylin precursor - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 22-Sep-1993 #sequence\_revision 26-May-1995 #text\_change 09-Jul-2004  
C:Accession: A55643; B46279  
R:Sciaky, D.; Kosiba, J.L.; Cohen, M.B.  
Genomics 24, 583-587, 1994  
A:Title: Genomic sequence of the murine guanylin gene.  
A:Reference number: A55643; MUID:95229161; PMID:7713512  
A:Accession: A55643  
A:Molecule type: DNA  
A:Residues: 1-116 <SCI>  
A:Cross-references: UNIPROT:P33680; GB:U60528; GB:U09741; NID:G1480667; PIDN:AA05758.1; R:de Sauvage, F.J.; Keshav, S.; Kuang, W.J.; Gillett, N.; Henzel, W.; Goeddel, D.V.  
Proc. Natl. Acad. Sci. U.S.A. 89, 9089-9093, 1992  
A:Title: Precursor structure, expression, and tissue distribution of human guanylin.  
A:Reference number: A46279; MUID:93028409; PMID:1409606  
A:Accession: B46279  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-116 <DEI>  
A:Cross-references: GB:M951175; NID:G309282; PIDN:AAA37758.1; PID:G309283  
A:Note: sequence extracted from NCBI backbone (NCBI:P:115379)  
C:Comment: Guanylin is an endogenous ligand for an intestine-specific receptor guanylyl n of the same receptor.  
C:Genetics:  
A:Introns: 25/3; 96/1  
C:Superfamily: guanylin  
C:Keywords: hormone; intestine  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-116/Product: guanylin #status predicted <MAT>

Query Match 26.5%; Score 153.5; DB 1; Length 116;  
Best Local Similarity 36.8%; Pred. No. 4.8e-09;  
Matches 42; Conservative 13; Mismatches 40; Indels 19; Gaps 5;

Qy 9 LLPGVAVVLLLLQSTQSVYIQGFVQLSMKSLDLEAQWAP-----SPRLQAO 60  
Db 11 LLGALAVLV-----EGTVQDGLSFPLESVKKLGKLEVOEPRLVSHKXKAPRL-LQ 62

Qy 61 SLLPAVC-HHPALPDLPQVCAQEAASSIFKTLRTIAND--DCELCVNVACTGC 111  
Db 63 PVAPQLCSSHSALEALRPVCEKPNAAEIIQLRLEIAQDPNTEICAYACTGC 116

RESULT 5  
JN0318  
guanylin precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 04-Dec-1992 #sequence\_revision 26-May-1995 #text\_change 09-Jul-2004  
C:Accession: JN0318; A43345; A38194; S25489  
R:Wiegand, R.C.; Kato, J.; Currie, M.G.  
Biochem. Biophys. Res. Commun. 185, 812-817, 1992  
A:Title: Rat guanylin cDNA: characterization of the precursor of an endogenous activator

[illegible]

**RESULT 8**

A47388  
serine/threonine protein kinase - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A47388  
R:Patriotis, C.; Makris, A.; Bear, S.E.; Tschlis, P.N.  
Proc. Natl. Acad. Sci. U.S.A. 90, 2251-2255, 1993  
A>Title: Tumor progression locus 2 (Tpl-2) encodes a protein kinase involved in the prog  
A/Reference number: A47388; MUID:93211939; PMID:7681591  
A:Accession: A47388  
A>Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-467 <PAR>  
A:Cross-references: UNIPROT:Q63562; GB:M94454; NID:g207082; PIDN:AAA42185.1; PID:g207083  
A:Experimental source: liver  
A>Note: sequence extracted from NCBI backbone (NCBIN:128133, NCBIPI:128134)  
F:136-388/Domain: protein kinase homology <KIN>

Query Match 13.0%; Score 75; DB 2; Length 467;  
Best Local Similarity 33.3%; Pred. No. 4.3;  
Matches 25; Conservative 15; Mismatches 25; Indels 10; Gaps 4;

QY 41 MKKLSLEAQWAPSPRLQAQSLLPAVCHHPAL--PQLQPVCASQEASSIFKTIRLIAND 98  
|::||::||::||::||::||::||::||::||::||::||::||::||::||::||:  
Db 361 MRELIEAALERNPHRPKAADL---KHEALNPREDPRCSLD-SALFDKRLLSRK 415  
|::||::||::||::||::||::||::||::||::||::||::||::||::||::||:  
QY 99 DCELCVNVA---CTG 110  
|::||::||::||  
Db 416 ELELPENIADSSCTG 430  
|::||::||::||

**RESULT 9**

A45082  
neurotrophic receptor rorl precursor - human  
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)  
C:Species: Homo sapiens (man)  
C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
C:Accession: A45082  
J:Maslakowski, P.; Carroll, R.D.  
R. Biol. Chem. 267, 26181-26190, 1992  
A>Title: A novel family of cell surface receptors with tyrosine kinase-like domain.  
A/Reference number: A45082; MUID:93100347; PMID:1334494  
A:Accession: A45082  
A:Molecule type: mRNA  
A:Residues: 1-937 <MAS>  
A:Cross-references: UNIPROT:Q01973; GB:M97675; NID:g337464; PIDN:AAA60275.1; PID:g337465  
A>Note: sequence extracted from NCBI backbone (NCBIPI:120916)  
C:Genetics:  
A:Gene: GDB:NTRKRI  
A:Cross-references: GDB:I36453  
A:Map position: gp21-gp21  
C:Superfamily: neurotrophic receptor ror; immunoglobulin homology; kringle homology; pro  
C:Keywords: ATP; glycoprotein; kringle; phosphotransferase; transmembrane protein; tyro  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-937/Product: neurotrophic receptor rorl #status predicted <MAT>  
F:72-133/Domain: immunoglobulin homology <IMM>  
F:72-133/Domain: kringle homology <KRG>  
F:313-391/Domain: transmembrane #status predicted <TM1>  
F:404-425/Domain: protein kinase homology <KIN>  
F:471-753/Domain: protein kinase homology <KIN>  
F:47,66,184,315/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 12.9%; Score 74.5; DB 2; Length 937;  
Best Local Similarity 24.1%; Pred. No. 10;  
Matches 26; Conservative 20; Mismatches 29; Indels 33; Gaps 6;

QY 12 GVAVLLLLLQSTQSVIVIQGFVRVQLSM-----KKLSLEAQWAPSPRLQAQ 60  
|::||::||::||::||::||::||::||::||::||::||::||::||::||::||:  
Db 175 GIACAFI---GNRTVMESLMHQEIENQITAAFTMTGTSSHLSDKCSQA----- 223  
|::||::||::||::||::||::||::||::||::||::||::||::||::||::||:  
QY 61 SLIPAVCHHPALPDQLQPVCAEQEASSIFKTIRLIANDDCELCNVAV 108  
|::||::||::||::||::||::||::||::||::||::||::||::||::||::||:

RESULT 12  
I50618  
C:fps proto oncogene - chicken  
C:Species: Gallus gallus (chicken)  
C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004  
C/Accession: I50618  
R:Huang, C.C.; Hammond, C.; Bishop, J.M.  
J. Mol. Biol. 181, 175-186, 1985  
A>Title: Nucleotide sequence and topography of chicken c-fps. Genesis of a retroviral on  
A:Reference number: I50618; MUID:85160839; PMID:3879969  
A/Accession: I50618  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-824 <HUA>  
A/Cross-references: UNIPROT:Q90778; EMBL:X02266; NID:G63203; PIDN:CAA26155.1; PID:g87104  
C:Genetics:  
A:Gene: c-fps  
A:Introns: 75/3; 133/3; 166/1; 227/2; 273/2; 313/2; 354/2; 416/3; 442/3; 512/3; 553/3; 5  
C:Superfamily: protein-tyrosine kinase fps; protein kinase homology; SH2 homology  
C:Keywords: ATP  
F:462-547/Domain: SH2 homology <SH2>  
F:561-823/Domain: protein kinase homology <KIN>  
F:569-577/Region: protein kinase ATP-binding motif

Query Match 12.8%; Score 74; DB 2; Length 824;  
Best Local Similarity 29.3%; Pred. No. 10;  
Matches 29; Conservative 13; Mismatches 31; Indels 26; Gaps 4;  
Mates 29; Conservative 13; Mismatches 31; Indels 26; Gaps 4;

Qy 9 LLPGVAVVLLLLQSTQSVYIQYQGF---RVQLSEMK-----KLSDLAEQWAPSPRLQAO 60  
Db 349 LSPGERVHLLKRGQLQEAQQLQGLVCAQAKLQQRDLMLAKLAEIGSERPP----- 401

Qy 61 SLLPAVCHHPALP--ODLPVCASQOEASSIFKTLRTTAN 97  
Db 402 -----PALPQEDRQSVCTDQERSGVTALETIKN 431

RESULT 13  
A48713  
serine/threonine-specific protein kinase cot, 58K form - human  
N/Alternate names: cot proto-oncogene, 58K form  
N/Contains: serine/threonine-specific protein kinase cot, 52K form  
C:Species: Homo sapiens (man)  
C>Date: 07-Apr-1994 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C/Accession: A48713; S31639  
R:Aoki, M.; Hamada, F.; Sugimoto, T.; Sumida, S.; Akiyama, T.; Toyoshima, K.  
J. Biol. Chem. 268, 22723-22732, 1993  
A>Title: The human cot proto-oncogene encodes two protein serine/threonine kinases with  
A:Reference number: A48713; MUID:94043034; PMID:8226782  
A/Accession: A48713  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-467 <AOK>  
A/Cross-references: UNIPROT:P41279  
A/Experimental source: TCO-4 cells  
A>Note: sequence inconsistent with nucleotide translation  
A:Note: sequence extracted from NCBI backbone (NCBIN:138970, NCBI:P:138971)  
A/Accession: B48713  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 30-467 <AO2>  
A/Experimental source: TCO-4 cells  
A:Note: sequence inconsistent with nucleotide translation  
A:Note: sequence extracted from NCBI backbone (NCBIN:138970, NCBI:P:138972)  
R:Chan, A.M.; Chedid, M.; Aaronson, S.A.; Mikki, T.; McGovern, E.S.  
Submitted to the EMBL Data Library, July 1992  
A/Description: A transforming gene isolated by expression cloning from Ewing's sarcoma c  
A:Reference number: S31639  
A/Accession: S31639  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-467 <CHA>

A/Cross-references: EMBL:Z14138; NID:g31244; PIDN:CAA78512.1; PID:g31245  
C:Keywords: alternative initiators; proto-oncogene  
F:136-388/Domain: protein kinase homology <KIN>

Query Match 12.4%; Score 72; DB 2; Length 467;  
Best Local Similarity 32.0%; Pred. No. 9;  
Matches 24; Conservative 16; Mismatches 25; Indels 10; Gaps 4;

Qy 41 MKKLSDLAEQWAPSPRLQAOQLPAPVCHHPALP---PQDLQPVCAEQEASSIFKTLRTTAND 98  
Db 361 MRELIEASLERPNVHRPRAADLL-----KHEALNPREDQPCQSLD-SALLERKLLSRK 415

Qy 99 DCELCVNVA---CTG 110  
Db 416 ELELPENIADSSCTG 430

RESULT 14  
S63453  
probable RNA helicase SUV3, ATP-dependent, precursor - yeast (Saccharomyces cerevisiae)  
N/Alternate names: protein LPB2w; protein YPL029w  
C:Species: Saccharomyces cerevisiae  
C>Date: 16-May-1996 #sequence\_revision 12-Jul-1996 #text\_change 09-Jul-2004  
C/Accession: S63453; A46190; S27462  
R:Wang, Y.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Hall, J.; Storms, R.K.; Vo  
submitted to the EMBL Data Library, September 1995  
A:Reference number: S63452  
A/Accession: S63453  
A:Molecule type: DNA  
A:Residues: 1-737 <WAN>  
A/Cross-references: UNIPROT:P32580; EMBL:U36624; NID:g1276642; PIDN:AAB68158.1; PID:g103  
R:Stepien, P.P.; Margossian, S.P.; Landman, D.; Butow, R.A.  
Proc. Natl. Acad. Sci. U.S.A. 89, 6813-6817, 1992  
A>Title: The yeast nuclear gene suv3 affecting mitochondrial post-transcriptional proces  
A:Reference number: A46190; MUID:92357722; PMID:1379722  
A/Accession: A46190  
A>Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-164, 'R', '166-597', 'DL', '600-615', 'A', '617-636', 'D', '638-663', 'D', '665-737' <STE>  
A/Cross-references: EMBL:M91167  
A/Note: sequence extracted from NCBI backbone (NCBI:P:111103)  
R:Stepien, P.P.; Margossian, S.P.; Landman, D.; Butow, R.A.  
submitted to the EMBL Data Library, April 1992  
A/Description: The yeast nuclear gene suv3 affecting mitochondrial post-transcriptional  
A:Reference number: S27462  
A/Accession: S27462  
A:Molecule type: DNA  
A:Residues: 1-164, 'R', '166-663', 'D', '665-737' <ST2>  
A/Cross-references: EMBL:M91167; NID:g172797; PIDN:AAA35135.1; PID:g172798  
C:Genetics:  
A:Gene: SGD:SUV3  
A/Cross-references: SGD:S0005950; MIPS:YPL029w  
A/Map position: 16L  
A/Genome: nuclear  
A/Function:  
A/Description: involved in mitochondrial translation and pre-mRNA splicing  
C:Keywords: ATP; mitochondrion; nucleotide binding; P-loop  
F:1-25/Domain: transit peptide (mitochondrion) #status predicted <TNP>  
F:26-737/Product: probable RNA helicase #status predicted <MAT>  
F:239-246/Region: nucleotide-binding motif A (P-loop)  
F:320-325/Region: nucleotide-binding motif B  
F:324-327/Region: DEAD/H motif #status atypical

Query Match 12.3%; Score 71; DB 2; Length 737;  
Best Local Similarity 26.8%; Pred. No. 19;  
Matches 21; Conservative 14; Mismatches 18; Indels 26; Gaps 3;

Qy 24 TQSVYIQYQGFVRVQLSEMK-----KLSDLAEQWAP-----SPRLQA 59  
Db 72 SNNVYLDSSFKQNLCKAMQFIYNDKLSLSDAKQVPIKNLAWLKLRDYIYQQLKDPKLOA 131

Qy 60 QSLLPVCH--HPALPQDL 76

Search completed: March 26, 2005, 16:49:40  
Job time : 23.708 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2005, 16:51:20 ; Search time 71.0357 Seconds  
(without alignments)  
129.758 Million cell updates/sec

Title: US-10-775-481A-3  
Perfect score: 18  
Sequence: 1 NTFYCCCLCCYPACAGCN 18

Scoring table:  Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0 1612378

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : UniProt\_03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID        | Description        |
|------------|-------|-------------|--------|--------------|--------------------|
| 1          | 18    | 100.0       | 18     | 1 HSTB_ECOLI | P01560 escherichia |
| 2          | 10    | 55.6        | 18     | 2 Q7M0J3     | Q7m0j3 citrobacter |
| 3          | 10    | 55.6        | 61     | 2 Q6VEG9     | Q6veg9 escherichia |
| 4          | 10    | 55.6        | 72     | 1 HST1_ECOLI | P01559 escherichia |
| 5          | 7     | 38.9        | 61     | 2 Q6VEG7     | Q6veg7 escherichia |
| 6          | 7     | 38.9        | 61     | 2 Q6VEG8     | Q6veg8 escherichia |
| 7          | 7     | 38.9        | 72     | 1 HST2_ECOLI | Q47185 escherichia |
| 8          | 7     | 38.9        | 72     | 1 HST3_ECOLI | P07965 escherichia |
| 9          | 7     | 38.9        | 120    | 2 Q858Z4     | Q858z4 bacterioph  |
| 10         | 6     | 33.3        | 66     | 1 HST_YERKR  | P31518 yersinia kr |
| 11         | 6     | 33.3        | 71     | 1 HST_YEREN  | P74977 yersinia en |
| 12         | 6     | 33.3        | 71     | 1 HSTB_YEREN | O50319 yersinia en |
| 13         | 6     | 33.3        | 72     | 1 HSTC_YEREN | Q8x4m8 escherichia |
| 14         | 6     | 33.3        | 90     | 2 Q8X4M8     | Q8x4m8 escherichia |
| 15         | 6     | 33.3        | 96     | 2 Q7Y2Z2     | O7y2z2 stx1 conver |
| 16         | 6     | 33.3        | 96     | 2 Q8T1L3     | Q8t1l3 bacterioph  |
| 17         | 6     | 33.3        | 130    | 2 Q8PNN9     | Q8pnn9 xanthomonas |
| 18         | 6     | 33.3        | 252    | 2 Q8CBH7     | Q8cbh7 mus musculu |
| 19         | 6     | 33.3        | 534    | 1 AMT1_CABEL | P54145 caenorhabdi |
| 20         | 6     | 33.3        | 556    | 2 Q8VUI3     | Q8vui3 desulfovibr |
| 21         | 6     | 33.3        | 607    | 2 Q8QVF2     | Q8qv2 mus musculu  |
| 22         | 6     | 33.3        | 609    | 2 Q7TMR1     | Q7tmr1 mus musculu |
| 23         | 6     | 33.3        | 738    | 2 Q8C6V9     | Q8c6v9 mus musculu |
| 24         | 6     | 33.3        | 876    | 2 Q8K0J1     | Q8k0j1 mus musculu |
| 25         | 6     | 33.3        | 1030   | 2 Q9LHT5     | Q9lht5 arabidopsis |
| 26         | 6     | 33.3        | 1210   | 2 Q8CJM2     | Q8cjm2 streptomyce |
| 27         | 6     | 33.3        | 1282   | 2 Q8C2F9     | Q8c2f9 mus musculu |
| 28         | 6     | 33.3        | 2483   | 1 MPRI_MOUSE | Q07113 mus musculu |
| 29         | 5     | 27.8        | 15     | 1 GUAN_DIDMA | P55936 didelphis m |
| 30         | 5     | 27.8        | 40     | 1 CXAC_CONGE | Q86rb2 conus geogr |
| 31         | 5     | 27.8        | 45     | 2 Q83AK0     | Q83ak0 coxiella bu |

|    |   |      |    |   |            |                    |
|----|---|------|----|---|------------|--------------------|
| 32 | 5 | 27.8 | 50 | 2 | Q83BF3     | Q83bf3 coxiella bu |
| 33 | 5 | 27.8 | 57 | 2 | Q83A72     | Q83a72 coxiella bu |
| 34 | 5 | 27.8 | 60 | 2 | Q96JP7     | Q96jp7 homo sapien |
| 35 | 5 | 27.8 | 63 | 2 | Q9STC3     | Q9stc3 elaeis guin |
| 36 | 5 | 27.8 | 65 | 2 | O22376     | O22376 oryza sativ |
| 37 | 5 | 27.8 | 65 | 2 | Q6ATL2     | Q6atl2 oryza sativ |
| 38 | 5 | 27.8 | 65 | 2 | Q62MM8     | Q62mm8 burkholderi |
| 39 | 5 | 27.8 | 66 | 2 | Q6C751     | Q6c751 yarrowia li |
| 40 | 5 | 27.8 | 66 | 2 | Q688G0     | Q688g0 oryza sativ |
| 41 | 5 | 27.8 | 66 | 2 | Q83AN4     | Q83an4 coxiella bu |
| 42 | 5 | 27.8 | 73 | 2 | Q83F74     | Q83f74 coxiella bu |
| 43 | 5 | 27.8 | 74 | 2 | Q82370     | Q82370 salmonella  |
| 44 | 5 | 27.8 | 75 | 1 | RC22_ARATH | RC2232 arabidopsis |
| 45 | 5 | 27.8 | 76 | 2 | Q83AQ6     | Q83aq6 coxiella bu |

ALIGNMENTS

RESULT 1  
HSTB\_ECOLI STANDARD; PRT; 18 AA.  
AC P01560;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Heat-stable enterotoxin ST-2 (ST-B).  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=O42:K86:H37 / 18D / ETEC;  
RX MEDLINE=81264141; PubMed=7021541;  
RA Chan S.-K., Giannella R.A.;  
RT "Amino acid sequence of heat-stable enterotoxin produced by  
Escherichia coli pathogenic for man.";  
RL J. Biol. Chem. 256:7744-7746(1981).  
[2]  
RP DISULFIDE BONDS.  
RX MEDLINE=87191003; PubMed=3552731; DOI=10.1016/0014-5793(87)80134-5;  
RA Shimomishi Y., Hidaka Y., Koizumi M., Hane M.,imoto S., Takeda T.,  
Miwatani T., Takeda Y.;  
RT "Mode of disulfide bond formation of a heat-stable enterotoxin (STh)  
produced by a human strain of enterotoxigenic Escherichia coli.";  
RL FEBS Lett. 215:165-170(1987).  
CC -1- FUNCTION: Toxin which activates the particulate form of guanylate  
cyclase and increases cyclic GMP levels within the host intestinal  
epithelial cells.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- SIMILARITY: Belongs to the heat-stable enterotoxin family.  
DR PIR; A01823; QHEC2.  
DR HSRP; P01559; LETN.  
DR InterPro; IPR001489; Enterotoxin HS.  
DR Pfam; PF02048; Enterotoxin HS; 1.  
DR PROSITE; PS00273; ENTEROTOXIN\_H\_STABLE; 1.  
KW Direct protein sequencing; Enterotoxin; Toxin.  
FT DISULFID 5 10  
FT DISULFID 6 14  
FT DISULFID 9 17  
SQ SEQUENCE 18 AA; 1978 MW; DDC975F49D60650 CRC64;

Query Match 100.0%; Score 18; DB 1; Length 18;  
Best Local Similarity 100.0%; Pred. No. 1.9e-14;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NTFYCCCLCCYPACAGCN 18  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1 NTFYCCCLCCYPACAGCN 18  
| | | | | | | | | | | | | | | | | | | | | |

RESULT 2

```

Q7MOU3
ID Q7MOU3 PRELIMINARY; PRT; 18 AA.
AC Q7MOU3
DT 01-MAR-2004 (TReMBLrel. 26, Created)
DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Heat-stable enterotoxin ST-1a.
OS Citrobacter freundii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Citrobacter.
OX NCBI_TaxID=546;
RN [1]
RP SEQUENCE.
RA MEDLINE=89108617; PubMed=2912902;
RA Guarino A., Giannella R., Thompson M.R.;
RT "Citrobacter freundii produces an 18-amino-acid heat-stable
RT enterotoxin identical to the 18-amino-acid Escherichia coli heat-
RT stable enterotoxin (ST 1a).";
RL Infect. Immun. 57:649-652(1989).
DR PIR: A60103; A60103.
DR HSP: P01559; IETN.
DR GO: GO:0005576; C:extracellular; IEA.
DR GO: GO:0009405; P:pathogenesis; IEA.
DR InterPro: IPR001489; Enterotoxin HS.
DR Pfam: PF02048; Enterotoxin HS; 1.
DR PROSITE: PS00273; ENTEROTOXIN_H_STABLE; 1.
SQ SEQUENCE 18 AA; 1978 MW; D4D975F49D60064F CRC64;

Query Match 55.6%; Score 10; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 7.2e-05;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTFYCCCLCC 10
Db 1 NTFYCCCLCC 10

RESULT 3
Q6VEG9 PRELIMINARY; PRT; 61 AA.
AC Q6VEG9
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Heat-stable enterotoxin ST-1a (Fragment).
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA Reischl U., Youssef M.T., Wolf H., Hyttia-Trees E., Strockbine N.A.;
RT "Real-time fluorescence PCR assays for detection and characterization
RT of heat-labile I and heat-stable I enterotoxin genes from
RT enterotoxigenic Escherichia coli.";
RL EMBL: AY342057; AAQ92974.1; -.
DR GO: GO:0005576; C:extracellular; IEA.
DR GO: GO:0009405; P:pathogenesis; IEA.
DR InterPro: IPR001489; Enterotoxin HS.
DR Pfam: PF02048; Enterotoxin HS; 1.
FT NON_TER 1
SQ SEQUENCE 61 AA; 6927 MW; 646D4AE2P8999957 CRC64;

Query Match 55.6%; Score 10; DB 2; Length 61;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTFYCCCLCC 10
Db 44 NTFYCCCLCC 53

```

```

RESULT 4
HSTL_ECOLI
ID HSTL_ECOLI STANDARD; PRT; 72 AA.
AC P01559; Q47653;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Heat-stable enterotoxin ST-1a/ST-P precursor.
GN Name=stl;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON=Tn1681;
RX MEDLINE=81054703; PubMed=6254008;
RA So M., McCarthy B.J.;
RT "Nucleotide sequence of the bacterial transposon Tn1681 encoding a
RT heat-stable (ST) toxin and its identification in enterotoxigenic
RT Escherichia coli strains.";
RL Proc. Natl. Acad. Sci. U.S.A. 77:4011-4015(1980).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=042.K86.H37 / 18D / ETEC;
RX MEDLINE=90368614; PubMed=2203756;
RA Dallas W.S.;
RT "The heat-stable toxin I gene from Escherichia coli 18D.";
RL J. Bacteriol. 172:5490-5493(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=85249571; PubMed=2990268;
RA Sekizaki T., Akashi H., Terakado N.;
RT "Nucleotide sequences of the genes for Escherichia coli heat-stable
RT enterotoxin I of bovine, avian, and porcine origins.";
RL Am. J. Vet. Res. 46:909-912(1985).
RN [4]
RP DISULFIDE BONDS.
RX MEDLINE=87191003; PubMed=3552731; DOI=10.1016/0014-5793(87)80134-5;
RA Shimonishi Y., Hidaka Y., Koizumi M., Hane M., Aimoto S., Takeda T.,
RA Miwatani T., Takeda Y.;
RT "Mode of disulfide bond formation of a heat-stable enterotoxin (STh)
RT produced by a human strain of enterotoxigenic Escherichia coli.";
RL FEBS Lett. 215:165-170(1987).
RN [5]
RP PROCESSING.
RX MEDLINE=90368584; PubMed=2203746;
RA Okamoto K., Takahara M.;
RT "Synthesis of Escherichia coli heat-stable enterotoxin STp as a pre-
RT pro form and role of the pro sequence in secretion.";
RL J. Bacteriol. 172:5260-5265(1990).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (0.89 ANGSTROMS) OF 59-71.
RX MEDLINE=94312375; PubMed=8038153;
RA Sato T., Ozaki H., Hata Y., Kitagawa Y., Kataube Y., Shimonishi Y.;
RT "Structural characteristics for biological activity of heat-stable
RT enterotoxin produced by enterotoxigenic Escherichia coli: X-ray
RT crystallography of weakly toxic and nontoxic analogs.";
RL Biochemistry 33:8641-8650(1994).
CC -!- FUNCTION: Toxin which activates the particulate form of guanylate
CC cyclase and increases cyclic GMP levels within the host intestinal
CC epithelial cells.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the heat-stable enterotoxin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```





```

CC DR EMBL; M18345; AAA23729.1; -
DR PIR; JS0292; QHECIB.
DR HSSP; P01559; IETN.
DR InterPro; IPR001489; Enterotoxin HS.
DR Pfam; PF02048; Enterotoxin HS; 1.
DR PROSITE; PS00273; ENTEROTOXIN_H_STABLE; 1.
KW Enterotoxin; Signal; Toxin.
FT SIGNAL 1 19 Potential.
FT PROPEP 20 53 By similarity.
FT PEPTIDE 54 72 Heat-stable enterotoxin A2.
FT DISULFID 59 64 By similarity.
FT DISULFID 60 68 By similarity.
FT DISULFID 63 71 By similarity.
FT DISULFID 63 71 By similarity.
SQ SEQUENCE 72 AA; 7895 MW; D87850306E06260 CRC64;

Query Match 38.9%; Score 7; DB 1; Length 72;
Best Local Similarity 100.0%; Pred. No. 0.73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 YCCELCC 10
DB 58 YCCELCC 64

RESULT 8
HST3_ECOLI STANDARD; PRT; 72 AA.
ID HST3_ECOLI
AC P07965; P26588;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Heat-stable enterotoxin A3/A4 precursor (ST3/ST44) (ST-IB) (ST-H).
GN Name:sta3; Synonym:sta4;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89202548; PubMed=3071819;
RA Stieglitz H., Cervantes L., Robledo R., Fonseca R., Covarrubias L.,
RA Bolivar F., Kupersztoch Y.M.;
RT "Cloning, sequencing, and expression in Ficol1-generated minicells of
RT an Escherichia coli heat-stable enterotoxin gene.";
RL Plasmid 20:42-53 (1988).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=153837-2;
RC STRAIN=83184648; PubMed=6341230;
RX MEDLINE=83184648; PubMed=6341230;
RA Moseley S.L., Hardy J.W., Huq M.I., Echeverria P., Falkow S.;
RT "Isolation and nucleotide sequence determination of a gene encoding a
RT heat-stable enterotoxin of Escherichia coli.";
RL Infect. Immun. 39:1167-1174 (1983).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=90273381; PubMed=2190361; DOI=10.1016/0041-0101(90)90085-L;
RA Zhou X., Shen L.P., Chi C.W.;
RT "Isolation and nucleotide sequence determination of a gene encoding a
RT heat-stable enterotoxin of Escherichia coli.";
RL Toxicon 28:453-456 (1990).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=89108616; PubMed=2643580;
RA Guzman-Verduzco L.M., Kupersztoch Y.M.;
RT "Rectification of two Escherichia coli heat-stable enterotoxin allele
RT sequences and lack of biological effect of changing the carboxy-
RT terminal tyrosine to histidine.";
RL Infect. Immun. 57:645-648 (1989).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=90034194; PubMed=2680769; DOI=10.1016/0378-1119(89)90182-0;
RA Dwarakanath P., Visweswariah S.S., Subrahmanyam Y.V.B.K., Shanthi G.,

```

---

```

RA Jagannatha H.M., Balganesht T.S.;
RT "Cloning and hyperexpression of a gene encoding the heat-stable toxin
RT of Escherichia coli.";
RL Gene 81:219-226 (1989).
RN [6]
RP SEQUENCE OF 54-72.
RX MEDLINE=83105138; PubMed=6759126;
RA Aimoto S., Takao T., Shimonishi Y., Hara S., Takeda T., Takeda Y.,
RA Miwatani T.;
RT "Amino-acid sequence of a heat-stable enterotoxin produced by human
RT enterotoxigenic Escherichia coli.";
RL Eur. J. Biochem. 129:257-263 (1982).
RN [7]
RP DISULFIDE BONDS.
RX MEDLINE=87191003; PubMed=3552731; DOI=10.1016/0014-5793(87)80134-5;
RA Shimonishi Y., Hidaka Y., Koizumi M., Hane M., Aimoto S., Takeda T.,
RA Miwatani T., Takeda Y.;
RT "Mode of disulfide bond formation of a heat-stable enterotoxin (STh)
RT produced by a human strain of enterotoxigenic Escherichia coli.";
RL FEBS Lett. 215:165-170 (1987).
RN [8]
RP PROCESSING.
RX MEDLINE=90251166; PubMed=2187146;
RA Rasheed J.K., Guzman-Verduzco L.M., Kupersztoch Y.M.;
RT "Two precursors of the heat-stable enterotoxin of Escherichia coli:
RT evidence of extracellular processing.";
RL Mol. Microbiol. 4:265-273 (1990).
CC -1- FUNCTION: Toxin which activates the particulate form of guanylate
CC cyclase and increases cyclic GMP levels within the host intestinal
CC epithelial cells.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the heat-stable enterotoxin family.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J03311; AAA24652.1; -
CC EMBL; M34916; AAA23990.1; -
CC EMBL; M18346; AAA23730.1; -
CC EMBL; M29255; AAA24686.1; -
CC PIR; JS0292; QHECIB.
CC PIR; JT0373; QHEC4.
CC HSSP; P01559; IETN.
CC InterPro; IPR001489; Enterotoxin HS.
CC Pfam; PF02048; Enterotoxin HS; 1.
CC PROSITE; PS00273; ENTEROTOXIN_H_STABLE; 1.
KW Direct protein sequencing; Enterotoxin; Signal; Toxin.
FT SIGNAL 1 19
FT PROPEP 20 53 Heat-stable enterotoxin A3/A4.
FT PEPTIDE 54 72
FT DISULFID 59 64
FT DISULFID 60 68
FT DISULFID 63 71
FT CONFLICT 19 19 A -> P (in Ref. 2).
SQ SEQUENCE 72 AA; 7909 MW; 1C5C9292BFCA6BA CRC64;

Query Match 38.9%; Score 7; DB 1; Length 72;
Best Local Similarity 100.0%; Pred. No. 0.73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 YCCELCC 10
DB 58 YCCELCC 64

RESULT 9
Q85824 PRELIMINARY; PRT; 120 AA.
ID Q85824

```

Q85824;  
 AC 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Gp7.  
 OS Bacteriophage phi-BT1.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;  
 OC Lambda-like viruses.  
 OX NCBI\_TaxID=225588;  
 RN [1]  
 RP MEDLINE=22803275; PubMed=12923110;  
 RX DOI=10.1128/JB.185.17.5320-5323.2003;  
 RA Gregory M.A., Till R., Smith M.C.M.;  
 RT "Integration Site for Streptomyces Phage {phi}BT1 and Development of  
 RT Site-Specific Integrating Vectors.";  
 RL J. Bacteriol. 185:5320-5323(2003).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Gregory M.A.;  
 RT "Characterisation and Evolution of Homologous Streptomyces  
 RT Bacteriophages";  
 RL Thesis (2000), Department of Genetics, University of Nottingham,  
 RL Nottingham, UNITED KINGDOM.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Smith M.C.M.;  
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ550940; CAD80131.1; -;  
 DR GO; GO:0004519; F:endonuclease activity; IEA.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR InterPro; IPR002711; HNH.  
 DR InterPro; IPR003615; HNH\_nuc.  
 DR Pfam; PF01844; HNH; 1.  
 DR SMART; SM00507; HNHc; 1.  
 SQ SEQUENCE 120 AA; 13827 MW; 2EAECL6389699723 CRC64;

Query Match 38.9%; Score 7; DB 2; Length 120;  
 Best Local Similarity 100.0%; Pred. No. 1;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 PACAGCN 18  
 Db 94 PACAGCN 100  
 |||||

RESULT 10  
 HST\_YERKR  
 ID HST\_YERKR STANDARD; PRT; 66 AA.  
 AC P31518;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Heat-stable enterotoxin precursor.  
 GN Name=yet;  
 OS Yersinia kristensenii.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Yersinia.  
 OX NCBI\_TaxID=28152;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ibrahim A., Liesack W., Stackebrandt E.;  
 RL Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Toxin which activates the particulate form of guanylate  
 CC cyclase and increases cyclic GMP levels within the host intestinal  
 CC epithelial cells (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Belongs to the heat-stable enterotoxin family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; X69218; CAA49152.1; -;  
 DR PIR; S31652; S31652.  
 DR HSSP; P01559; LETN.  
 DR InterPro; IPR001489; Enterotoxin HS.  
 DR Pfam; PF02048; Enterotoxin HS; 1.  
 DR PROSITE; PS00273; ENTEROTOXIN\_H\_STABLE; 1.  
 KW Enterotoxin; Signal.  
 FT SIGNAL 1 19 Potential.  
 FT PROPEP 20 50  
 FT CHAIN 51 66 Heat-stable enterotoxin (By similarity).  
 FT DISULFID 54 59 By similarity.  
 FT DISULFID 55 63 By similarity.  
 FT DISULFID 58 66 By similarity.  
 SQ SEQUENCE 66 AA; 7068 MW; 27BE7006675CC075 CRC64;

Query Match 33.3%; Score 6; DB 1; Length 66;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 PACAGC 17  
 Db 61 PACAGC 66  
 |||||

RESULT 11  
 HSTA\_YEREN  
 ID HSTA\_YEREN STANDARD; PRT; 71 AA.  
 AC P07593;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Heat-stable enterotoxin A precursor (YST-A).  
 GN Name=yetA; Synonyms=yet;  
 OS Yersinia enterocolitica.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Yersinia.  
 OX NCBI\_TaxID=630;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=Serotype O:8;  
 RX MEDLINE=93051243; PubMed=1427005; DOI=10.1016/0378-1097(92)90364-T;  
 RA Ibrahim A., Liesack W., Pike S., Stackebrandt E.;  
 RT "The polymerase chain reaction: an epidemiological tool to  
 RT differentiate between two clusters of pathogenic Yersinia  
 RT enterocolitica strains.";  
 RL FEMS Microbiol. Lett. 76:63-66(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=W1024 / Serotype O:9;  
 RX MEDLINE=90354067; PubMed=2201642;  
 RA Delor I., Kaeckenbeeck A., Wauters G., Cornelis G.R.;  
 RT "Nucleotide sequence of yet, the Yersinia enterocolitica gene encoding  
 RT the heat-stable enterotoxin, and prevalence of the gene among  
 RT pathogenic and nonpathogenic yersiniae.";  
 RL Infect. Immun. 58:2983-2988(1990).  
 RN [3]  
 RP SEQUENCE OF 54-71.  
 RA MEDLINE=86004705; PubMed=4043080;  
 RA Takao T., Tomimaga N., Yoshimura S., Shimonishi Y., Hara S., Inoue T.,  
 RA Miyama A.;  
 RT "Isolation, primary structure and synthesis of heat-stable enterotoxin  
 RT produced by Yersinia enterocolitica.";  
 RL Eur. J. Biochem. 152:199-206(1985).  
 CC -1- FUNCTION: Toxin which activates the particulate form of guanylate  
 CC cyclase and increases cyclic GMP levels within the host intestinal  
 CC epithelial cells.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- INDUCTION: In cultured cells, expressed only at temperatures <30

degrees Celsius. Under conditions of high osmolarity and alkaline pH (as it is the case in the host's intestine), it is expressed at 37 degrees Celsius.

-1- SIMILARITY: Belongs to the heat-stable enterotoxin family.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; X65999; CAA46801.1; -;  
EMBL; U09235; AAA18472.1; -;  
PIR; S25659; S25659.  
HSSP; P01559; IETN.  
InterPro; IPR001489; Enterotoxin HS.  
Pfam; PF02048; Enterotoxin HS; 1.  
PROSITE; PS00273; ENTEROTOXIN\_H STABLE; 1.  
Direct protein sequencing; Enterotoxin; Signal; Toxin.  
KW SIGNAL 1 19 Potential.  
FT PROPEP 20 53  
FT CHAIN 54 71 Heat-stable enterotoxin A.  
FT DISULFID 59 64 By similarity.  
FT DISULFID 60 68 By similarity.  
FT DISULFID 63 71 By similarity.  
FT VARIANT 48 48 L->S.  
FT SEQUENCE 71 AA; 7494 MW; 022F99B3800C861B CRC64;

Query Match 33.3%; Score 6; DB 1; Length 71;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 PACAGC 17  
Db 66 PACAGC 71

RESULT 12  
HSTB\_YEREN STANDARD; PRT; 71 AA.  
AC P74977;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Heat-stable enterotoxin B precursor (Y-STB).  
GN Names:yStB;  
OS Yersinia enterocolitica.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Yersinia.  
OX NCBI\_TaxID=630;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC STRAIN=84-50 / Serotype O:5;  
RC MEDLINE=98008784; PubMed=9344780; DOI=10.1006/mpat.1997.0146;  
RA Ramamurthy T., Yoshino K., Huang X., Nair G.B., Carniel E.,  
RA Maruyama T., Fukushima H., Takeda T.;  
RT "The novel heat-stable enterotoxin subtype gene (yStB) of Yersinia enterocolitica: nucleotide sequence and distribution of the yst genes.";  
RL Microb. Pathog. 23:189-200 (1997).  
CC -1- FUNCTION: Toxin which activates the particulate form of guanylate cyclase and increases cyclic GMP levels within the host intestinal epithelial cells. Could play an important role in pathogenesis.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- SIMILARITY: Belongs to the heat-stable enterotoxin family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; D88145; BAA13544.1; -;  
HSSP; P01559; IETN.  
InterPro; IPR001489; Enterotoxin HS.  
Pfam; PF02048; Enterotoxin HS; 1.  
PROSITE; PS00273; ENTEROTOXIN\_H STABLE; 1.  
Direct protein sequencing; Enterotoxin; Signal; Toxin.  
KW SIGNAL 1 19 Potential.  
FT PROPEP 20 52  
FT CHAIN 53 71 Heat-stable enterotoxin B.  
FT DISULFID 59 64 By similarity.  
FT DISULFID 60 68 By similarity.  
FT DISULFID 63 71 By similarity.  
FT SEQUENCE 71 AA; 7670 MW; ED6E9F61ACDD4F50 CRC64;

Query Match 33.3%; Score 6; DB 1; Length 71;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 PACAGC 17  
Db 66 PACAGC 71

RESULT 13  
HSTC\_YEREN STANDARD; PRT; 72 AA.  
AC O50319;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Heat-stable enterotoxin C precursor (Y-STC).  
GN Name:yStC;  
OS Yersinia enterocolitica.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Yersinia.  
OX NCBI\_TaxID=630;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=86-11;  
RC MEDLINE=97202479; PubMed=9049998; DOI=10.1006/mpat.1996.0094;  
RA Huang X., Yoshino K., Nakao H., Takeda T.;  
RT "Nucleotide sequence of a gene encoding the novel Yersinia enterocolitica heat-stable enterotoxin that includes a pro-region-like sequence in its mature toxin molecule.";  
RL Microb. Pathog. 22:89-97 (1997).  
RN [2]  
RP SEQUENCE OF 20-72.  
RC STRAIN=Serotype O:3;  
RC MEDLINE=95246844; PubMed=7729521; DOI=10.1016/0014-5793(95)00267-D;  
RA Yoshino K., Takao T., Huang X., Murata H., Nakao H., Takeda T.,  
RA Shimonishi Y.;  
RT "Characterization of a highly toxic, large molecular size heat-stable enterotoxin produced by a clinical isolate of Yersinia enterocolitica.";  
RL FEBS Lett. 362:319-322 (1995).  
CC -1- FUNCTION: Toxin which activates the particulate form of guanylate cyclase and increases cyclic GMP levels within the host intestinal epithelial cells. Highly toxic.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- SIMILARITY: Belongs to the heat-stable enterotoxin family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

DR HSP; P01559; IETN.
DR InterPro; IPR001489; Enterotoxin HS.
DR Pfam; PF02048; Enterotoxin HS; 1.
DR PROSITE; PS00273; ENTEROTOXIN_H STABLE; 1.
KW Direct protein sequencing; Enterotoxin; Signal; Toxin.
FT SIGNAL 1 19
FT CHAIN 20 72 Heat-stable enterotoxin C.
FT DISULFID 60 65 By similarity.
FT DISULFID 61 69 By similarity.
FT DISULFID 64 72 By similarity.
SQ SEQUENCE 72 AA; 7639 MW; 7C0D83893C2F981D CRC64;

Query Match 33.3%; Score 6; DB 1; Length 72;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 CCELC 10
Db 60 CCELC 65

RESULT 14
Q8X4M8 PRELIMINARY; PRT; 90 AA.
AC Q8X4M8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative lysis protein S of prophage CP-933V.
GN OrderedLocusNames=z3340;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posaf G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimailanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 408:529-533 (2001).
DR EMBL; AE005442; AAG57224.1; -.
DR FIR; D85845; D85845.
DR Pfam; PF04971; Lysis_S; 1.
KW Complete proteome.
SQ SEQUENCE 90 AA; 10060 MW; 680A0C48B94B2AF CRC64;

Query Match 33.3%; Score 6; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 FYCCEL 8
Db 7 FYCCEL 12

RESULT 15
Q7Y222 PRELIMINARY; PRT; 96 AA.
AC Q7Y222;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Lysis protein.
OS Stx1 converting bacteriophage.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX NCBI_TaxID=194948;

```

---

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Stx1 phage;
RX MEDLINE=22697399; PubMed=12813092;
RX DOI=10.1128/JB.185.13.3966-3971.2003;
RA Sato T., Shimizu T., Watarai M., Kobayashi M., Kano S., Hamabata T.,
RA Takeda Y., Yamasaki S.;
RT "Genome analysis of a novel Shiga toxin 1 (Stx1)-converting phage
RT which is closely related to Stx2-converting phages but not to other
RT Stx1-converting phages.";
RL J. Bacteriol. 185:3966-3971 (2003).
DR EMBL; AP005153; BAC77969.1; -.
DR InterPro; IPR007054; Lysis_S; 1.
DR Pfam; PF04971; Lysis_S; 1.
SQ SEQUENCE 96 AA; 10684 MW; CF1562C30DA56B19 CRC64;

Query Match 33.3%; Score 6; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 FYCCEL 8
Db 13 FYCCEL 18

Search completed: March 26, 2005, 17:24:51
Job time : 72.0357 secs

```

***This Page Blank (uspto)***

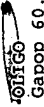
GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2005, 16:49:45 ; Search time 84.4821 Seconds  
(without alignments)  
86.982 Million cell updates/sec

Title: US-10-775-481A-5

Perfect score: 19  
Sequence: 1 NSSNYCCELCNACNGCY 19

Scoring table:   
Gapop 60.0 , Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003bs:\*
- 8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 19    | 100.0       | 19     | 2     | AAR85948    |
| 2          | 19    | 100.0       | 19     | 2     | AAY40510    |
| 3          | 19    | 100.0       | 19     | 8     | ADR45825    |
| 4          | 15    | 78.9        | 18     | 2     | AAR85970    |
| 5          | 15    | 78.9        | 18     | 2     | AAY40532    |
| 6          | 15    | 78.9        | 18     | 8     | ADR45846    |
| 7          | 15    | 78.9        | 19     | 2     | AAR95923    |
| 8          | 15    | 78.9        | 19     | 2     | RAY02386    |
| 9          | 15    | 78.9        | 19     | 2     | AAY29608    |
| 10         | 15    | 78.9        | 19     | 2     | RAY06972    |
| 11         | 15    | 78.9        | 19     | 2     | RAY02398    |
| 12         | 15    | 78.9        | 19     | 4     | AAM51879    |
| 13         | 15    | 78.9        | 19     | 6     | RAO16204    |
| 14         | 15    | 78.9        | 19     | 6     | ABG74825    |
| 15         | 15    | 78.9        | 19     | 7     | ADC14120    |
| 16         | 15    | 78.9        | 19     | 7     | ADC14123    |
| 17         | 15    | 78.9        | 19     | 7     | ADC14121    |
| 18         | 15    | 78.9        | 19     | 8     | ADR48340    |
| 19         | 15    | 78.9        | 19     | 8     | ADR48360    |
| 20         | 15    | 78.9        | 19     | 8     | ADR48359    |
| 21         | 15    | 78.9        | 19     | 8     | ADR48398    |
| 22         | 15    | 78.9        | 19     | 8     | ADR48329    |
| 23         | 15    | 78.9        | 19     | 8     | ADR48404    |
| 24         | 15    | 78.9        | 21     | 8     | ADR48372    |
| 25         | 15    | 78.9        | 21     | 8     | ADR48413    |

|    |    |      |    |   |          |                    |
|----|----|------|----|---|----------|--------------------|
| 26 | 15 | 78.9 | 21 | 8 | ADR48373 | Adr48373 Analgesic |
| 27 | 15 | 78.9 | 21 | 8 | ADR48414 | Adr48414 GC-C acti |
| 28 | 15 | 78.9 | 72 | 2 | AAR20352 | Aar20352 Sequence  |
| 29 | 15 | 78.9 | 72 | 2 | AAR71975 | Aar71975 E. coli s |
| 30 | 15 | 78.9 | 72 | 8 | ADR48356 | Adr48356 Immature  |
| 31 | 14 | 73.7 | 17 | 2 | AAR85971 | Aar85971 ST Ib rec |
| 32 | 14 | 73.7 | 17 | 2 | AAY40533 | Aay40533 ST recept |
| 33 | 14 | 73.7 | 17 | 8 | ADR45847 | Adr45847 Amino aci |
| 34 | 14 | 73.7 | 18 | 2 | AAR85976 | Aar85976 ST Ib rec |
| 35 | 14 | 73.7 | 18 | 2 | AAY40538 | Aay40538 ST recept |
| 36 | 14 | 73.7 | 18 | 8 | ADR45852 | Adr45852 Amino aci |
| 37 | 14 | 73.7 | 19 | 4 | AAB83216 | Aab83216 Escherich |
| 38 | 13 | 68.4 | 16 | 2 | AAR85972 | Aar85972 ST Ib rec |
| 39 | 13 | 68.4 | 16 | 2 | AAY40534 | Aay40534 ST recept |
| 40 | 13 | 68.4 | 16 | 8 | ADR45848 | Adr45848 Amino aci |
| 41 | 13 | 68.4 | 17 | 2 | AAR85977 | Aar85977 ST Ib rec |
| 42 | 13 | 68.4 | 17 | 2 | AAY40539 | Aay40539 ST recept |
| 43 | 13 | 68.4 | 17 | 8 | ADR45853 | Adr45853 Amino aci |
| 44 | 12 | 63.2 | 15 | 2 | AAR85973 | Aar85973 ST Ib rec |
| 45 | 12 | 63.2 | 15 | 2 | AAY40535 | Aay40535 ST recept |

#### ALIGNMENTS

##### RESULT 1

AAR85948  
ID AAR85948 standard; peptide; 19 AA.

AC AAR85948;

XX  
DT 19-JAN-1996 (first entry)

DE ST Ib receptor ligand portion.

XX  
KW ST; receptor; ligand; heat-stable toxin; conjugate; imaging; cancer;  
colorectal; metastasis.

XX  
OS Escherichia coli.

XX  
PN WO9511694-A1.

XX  
PD 04-MAY-1995.

XX  
PF 26-OCT-1994; 94WO-US012232.

XX  
PR 26-OCT-1993; 93US-00141892.

XX  
PR 13-SEP-1994; 94US-00305056.

XX  
(UYJE-) UNIV JEFFERSON THOMAS.

XX  
Waldman SA;

XX  
DR WPI; 1995-1786646/23.

XX  
PT Conjugated cpds. which specifically bind to colorectal cancer cells -  
comprise heat-stable toxin receptor binding moiety and active moiety  
which may be a therapeutic agent or a radioactive agent.

XX  
PS Claim 3; Page 116; 133pp; English.

XX  
CC New conjugated compounds are provided which consist of (1) an ST receptor  
binding moiety and (2) an active moiety which is a radio- stable agent.  
CC 'ST' refers to a heat stable toxin produced by E.coli and other  
CC organisms. Especially the radiostable agent is a therapeutic agent (e.g.  
CC methotrexate, doxorubicin, mitomycin), imaging agent (e.g. radioactive  
CC isotope) or nucleic acid; and the compound is used for the detection,  
CC imaging or treatment of colorectal tumours, particularly metastasised  
CC tumours. The present sequence is a specific example of an ST receptor  
CC binding peptide which can be used in the conjugate

XX  
SQ Sequence 19 AA;

Query Match 100.0%; Score 19; DB 2; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-14;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPACNGCY 19  
 Db 1 NSSNYCCCLCCNPACNGCY 19

RESULT 2  
 AAY40510  
 ID AAY40510 standard; peptide; 19 AA.  
 XX AC AAY40510;  
 XX DT 03-DEC-1999 (first entry)  
 XX DE ST receptor peptide ST 1b.  
 XX KW Conjugated compound; ST receptor; heat-stable; toxin; colorectal cancer.  
 XX OS Escherichia coli.  
 XX PN US5962220-A.  
 XX PD 05-OCT-1999.  
 XX PF 06-JUN-1995; 95US-00467920.  
 XX PR 26-OCT-1993; 93US-00141892.  
 XX PA (UYJE-) UNIV JEFFERSON THOMAS.  
 XX PI Waldman SA;  
 XX DR WPI; 1999-571264/48.  
 XX DR N-PSDB; AAZ07541.  
 XX PT Conjugated compound comprising a receptor moiety and active moiety,  
 XX PT useful for the treatment/prevention of colorectal cancer.  
 XX PS Claim 1; Col 21-22; 23pp; English.  
 XX SQ The invention relates to a conjugated compound that comprises a ST (heat-stable toxin) receptor moiety selected from one of the sequences shown in CC  
 CC AAY40508-Y40559 and an active moiety (antisense molecule). The compound CC  
 CC is useful for the treatment of colorectal cancer. The use of the CC  
 CC conjugated compound is advantageous compared to antibodies since it binds CC  
 CC specifically to colorectal cells and has no toxic effect on normal cells. CC  
 CC The present sequence represents the amino acid sequence of a ST receptor CC  
 CC peptide ST 1b  
 XX SQ Sequence 19 AA;

Query Match 100.0%; Score 19; DB 2; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-14;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPACNGCY 19  
 Db 1 NSSNYCCCLCCNPACNGCY 19

RESULT 3  
 ADR45825  
 ID ADR45825 standard; peptide; 19 AA.  
 XX AC ADR45825;  
 XX DT 18-NOV-2004 (first entry)  
 XX DE Amino acid sequence of heat stable toxin ST 1b.  
 XX PD

Query Match 100.0%; Score 19; DB 8; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-14;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPACNGCY 19  
 Db 1 NSSNYCCCLCCNPACNGCY 19

RESULT 4  
 AAR85970  
 ID AAR85970 standard; peptide; 18 AA.  
 XX AC AAR85970;  
 XX DT 19-JAN-1996 (first entry)  
 XX DE ST 1b receptor ligand portion.  
 XX KW ST; receptor; ligand; heat-stable toxin; conjugate; imaging; cancer;  
 XX KW colorectal; metastasis.  
 XX OS Escherichia coli.  
 XX PN W09511694-A1.  
 XX PD 04-MAY-1995.  
 XX PD

KW heat stable toxin; ST; ST receptor; colorectal cancer; gastric cancer;  
 KW oesophageal cancer; colorectal cell; gastric cell; oesophageal cell;  
 KW ST 1b.  
 XX Unidentified.  
 XX OS WO2004071436-A2.  
 XX PN 26-AUG-2004.  
 XX PD 10-FEB-2004; 2004WO-US003765.  
 XX PF 10-FEB-2003; 2003US-0446730P.  
 XX PR (UYJE-) UNIV JEFFERSON THOMAS.  
 XX PA Waldman SA, Pitari GM, Park J, Schulz S, Wolfe HR;  
 XX PI WPI; 2004-615913/59.  
 XX DR N-PSDB; ADR45824.  
 XX DR Increasing ST receptor molecules on the surface of a colorectal, gastric  
 XX PT or esophageal cancer cell to treat these cancers comprises administering  
 XX PT ST receptor ligand molecules that bind to ST receptors on the surface of  
 XX PT the cancer cell.  
 XX PS Claim 6; SEQ ID NO 5; 97pp; English.  
 XX CC The specification describes a method for increasing the number of heat  
 CC stable toxin (ST) receptor molecules on the surface of a metastasised  
 CC colorectal cancer cell. The method comprises administering, by continuous  
 CC infusion, at least 0.1 nM of an ST receptor ligand per 10 kg bodyweight  
 CC per hour for at least 6 hours, where ST receptor ligand molecules bind to  
 CC ST receptors on the surface of the cancer cell in the individual and the  
 CC number of ST receptor molecules on the surface of the cancer cell is  
 CC increased. Therapeutic compositions comprising components which target ST  
 CC receptors can then be used to inhibit proliferation of the colorectal,  
 CC gastric and oesophageal cells. This method may be used for treating  
 CC individuals that have diseases that affect colorectal, gastric and  
 CC oesophageal cells, including colorectal, gastric or oesophageal cancers.  
 CC The present sequence represents a ST, designated ST 1b, which is used as  
 CC the ST receptor ligand in the method of the invention.  
 XX SQ Sequence 19 AA;



PF 26-OCT-1994; 94WO-US012232.  
 XX  
 PR 26-OCT-1993; 93US-00141892.  
 PR 13-SEP-1994; 94US-00305056.  
 XX  
 PA (UYJE-) UNIV JEFFERSON THOMAS.  
 XX  
 PI Waldman SA;  
 XX  
 DR WPI; 1995-178646/23.  
 XX  
 PT Conjugated cpds. which specifically bind to colorectal cancer cells -  
 PT comprise heat-stable toxin receptor binding moiety and active moiety  
 PT which may be a therapeutic agent or a radioactive agent.  
 XX  
 PS Claim 3; Page 120; 133pp; English.  
 XX  
 CC New conjugated compounds are provided which consist of (1) an ST receptor  
 CC binding moiety and (2) an active moiety which is a radio-stable agent.  
 CC 'ST' refers to a heat stable toxin produced by E.coli and other  
 CC organisms. Especially the radiostable agent is a therapeutic agent (e.g.  
 CC methotrexate, doxorubicin, mitomycin), imaging agent (e.g. radioactive  
 CC isotope) or nucleic acid; and the compound is used for the detection,  
 CC imaging or treatment of colorectal tumours, particularly metastasised  
 CC tumours. The present sequence is a specific example of an ST receptor  
 CC binding peptide which can be used in the conjugate  
 XX  
 SQ Sequence 18 AA;  
 Query Match 78.9%; Score 15; DB 2; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-09;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NSSNYCCCLCCNPAC 15  
 DB 1 NSSNYCCCLCCNPAC 15  
 RESULT 5  
 AAY40532  
 ID AAY40532 standard; peptide; 18 AA.  
 XX  
 AC AAY40532;  
 XX  
 DT 03-DEC-1999 (first entry)  
 XX  
 DE ST receptor binding peptide.  
 XX  
 KW Conjugated compound; ST receptor; heat-stable; toxin; colorectal cancer.  
 XX  
 OS Escherichia coli.  
 XX  
 PN US962220-A.  
 XX  
 PD 05-OCT-1999.  
 XX  
 PF 06-JUN-1995; 95US-00467920.  
 XX  
 PR 26-OCT-1993; 93US-00141892.  
 XX  
 PA (UYJE-) UNIV JEFFERSON THOMAS.  
 XX  
 PI Waldman SA;  
 XX  
 DR WPI; 1999-571264/48.  
 XX  
 PT Conjugated compound comprising a receptor moiety and active moiety,  
 PT useful for the treatment/prevention of colorectal cancer.  
 XX  
 PS Claim 1; Col 29-30; 23pp; English.  
 XX  
 CC The invention relates to a conjugated compound that comprises a ST (heat-  
 CC stable toxin) receptor moiety selected from one of the sequences shown in  
 CC

CC AAY40508-Y40559 and an active moiety (antisense molecule). The compound  
 CC is useful for the treatment of colorectal cancer. The use of the  
 CC conjugated compound is advantageous compared to antibodies since it binds  
 CC specifically to colorectal cells and has no toxic effect on normal cells  
 XX  
 SQ Sequence 18 AA;  
 Query Match 78.9%; Score 15; DB 2; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-09;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NSSNYCCCLCCNPAC 15  
 DB 1 NSSNYCCCLCCNPAC 15  
 RESULT 6  
 ADR45846  
 ID ADR45846 standard; peptide; 18 AA.  
 XX  
 AC ADR45846;  
 XX  
 DT 18-NOV-2004 (first entry)  
 XX  
 DE Amino acid sequence of heat stable toxin ST 1b fragment.  
 XX  
 KW heat stable toxin; ST; ST receptor; colorectal cancer; gastric cancer;  
 KW oesophageal cancer; colorectal cell; gastric cell; oesophageal cell;  
 KW ST 1b.  
 XX  
 OS Unidentified.  
 XX  
 PN WO2004071436-A2.  
 XX  
 PD 26-AUG-2004.  
 XX  
 PF 10-FEB-2004; 2004WO-US003765.  
 XX  
 PR 10-FEB-2003; 2003US-0446730P.  
 XX  
 PA (UYJE-) UNIV JEFFERSON THOMAS.  
 XX  
 PI Waldman SA, Pitari GM, Park J, Schulz S, Wolfe HR;  
 XX  
 DR WPI; 2004-615913/59.  
 XX  
 DE Increasing ST receptor molecules on the surface of a colorectal, gastric  
 DE or oesophageal cancer cell to treat these cancers comprises administering  
 DE ST receptor ligand molecules that bind to ST receptors on the surface of  
 DE the cancer cell.  
 XX  
 PS Claim 6; SEQ ID NO 27; 97pp; English.  
 XX  
 CC The specification describes a method for increasing the number of heat  
 CC stable toxin (ST) receptor molecules on the surface of a metastasised  
 CC colorectal cancer cell. The method comprises administering, by continuous  
 CC infusion, at least 0.1 nM of an ST receptor ligand per 10 kg bodyweight  
 CC per hour for at least 6 hours, where ST receptor ligand molecules bind to  
 CC ST receptors on the surface of the cancer cell in the individual and the  
 CC number of ST receptor molecules on the surface of the cancer cell is  
 CC increased. Therapeutic compositions comprising components which target ST  
 CC receptors can then be used to inhibit proliferation of the colorectal,  
 CC gastric and oesophageal cells. This method may be used for treating  
 CC individuals that have diseases that affect colorectal, gastric and  
 CC oesophageal cells, including colorectal, gastric or oesophageal cancers.  
 CC The present sequence represents a fragment of a ST designated ST 1b (see  
 CC ADR45825), which is used as the ST receptor ligand in the method of the  
 CC invention.  
 XX  
 SQ Sequence 18 AA;  
 Query Match 78.9%; Score 15; DB 8; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-09;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPAC 15  
 Db 1 NSSNYCCCLCCNPAC 15

RESULT 7  
 AAR95923  
 ID AAR95923 standard; peptide; 19 AA.  
 AC AAR95923;  
 DT 14-JAN-1997 (first entry)  
 DE Enterotoxigenic E. coli STa protein antigen.  
 KW Mucosal binding composition; mucosal binding polypeptide; antigen;  
 KW non-viral pathogen; sexually transmitted disease; administration;  
 KW vaginal; rectal; oral; immune response; secretory immunity; mucous;  
 KW enterotoxigenic; STa protein.  
 XX Escherichia coli.  
 OS WO9616178-A1.  
 PN 30-MAY-1996.  
 PD 17-NOV-1995; 95WO-GB002708.  
 XX 17-NOV-1994; 94US-00342241.  
 XX (LEBE/) LEBENS M R.  
 PA (HOLM/) HOLMGREN J R.  
 XX Lebens MR, Holmgren JR;  
 PI WPI; 1996-268614/27.  
 XX Mucosal binding compositions for generating mucosal immune response -  
 PT comprises mucosal binding peptide, pref. derived from cholera toxin, and  
 PT an antigen, e.g. derived from E. coli, HIV, etc.  
 XX Claim 36; Page 44; 65pp; English.  
 CC A novel mucosal binding compsn. (MBC) comprises a mucosal binding  
 CC polypeptide linked to at least 1 antigen from a non-viral pathogen, which  
 CC causes a sexually transmitted disease (STD), e.g. the present  
 CC enterotoxigenic E. coli STa protein antigen. The MBC, which is  
 CC administered vaginally, rectally or orally, generates a mucosal immune  
 CC response against the viral STD by allowing for the prodn. of high levels  
 CC of secretory immunity, which forms the 1st line of defence against the  
 CC majority of STD  
 XX Sequence 19 AA;

Query Match 78.9%; Score 15; DB 2; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-09;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPAC 15  
 Db 1 NSSNYCCCLCCNPAC 15

RESULT 8  
 AAY02386  
 ID AAY02386 standard; peptide; 19 AA.  
 AC AAY02386;  
 DT 09-JUL-1999 (first entry)  
 DE Enterotoxigenic E. coli STa protein antigen.  
 KW Mucosal binding composition; mucosal binding polypeptide; antigen;  
 KW non-viral pathogen; sexually transmitted disease; administration;  
 KW vaginal; rectal; oral; immune response; secretory immunity; mucous;  
 KW enterotoxigenic; STa protein.  
 XX Escherichia coli.  
 OS WO9616178-A1.  
 PN 30-MAY-1996.  
 PD 17-NOV-1995; 95WO-GB002708.  
 XX 17-NOV-1994; 94US-00342241.  
 XX (LEBE/) LEBENS M R.  
 PA (HOLM/) HOLMGREN J R.  
 XX Lebens MR, Holmgren JR;  
 PI WPI; 1996-268614/27.  
 XX Mucosal binding compositions for generating mucosal immune response -  
 PT comprises mucosal binding peptide, pref. derived from cholera toxin, and  
 PT an antigen, e.g. derived from E. coli, HIV, etc.  
 XX Claim 36; Page 44; 65pp; English.  
 CC A novel mucosal binding compsn. (MBC) comprises a mucosal binding  
 CC polypeptide linked to at least 1 antigen from a non-viral pathogen, which  
 CC causes a sexually transmitted disease (STD), e.g. the present  
 CC enterotoxigenic E. coli STa protein antigen. The MBC, which is  
 CC administered vaginally, rectally or orally, generates a mucosal immune  
 CC response against the viral STD by allowing for the prodn. of high levels  
 CC of secretory immunity, which forms the 1st line of defence against the  
 CC majority of STD  
 XX Sequence 19 AA;

Query Match 78.9%; Score 15; DB 2; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-09;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPAC 15  
 Db 1 NSSNYCCCLCCNPAC 15

RESULT 9  
 AAY29608  
 ID AAY29608 standard; peptide; 19 AA.  
 AC AAY29608;  
 DT 15-OCT-1999 (first entry)  
 DE Escherichia coli heat stable ST enterotoxin STH.  
 KW Heat stable ST enterotoxin; immunoreagent; radiological therapy;  
 KW diagnosis; ST receptor binding moiety; macrocyclic complexing agent;  
 KW tumour; infectious diarrhoeal disease; diarrhoea.  
 XX Escherichia coli.  
 OS WO9939748-A1.  
 PN 12-AUG-1999.  
 PD 08-FEB-1999; 99WO-GB000396.  
 XX

DE Heat stable ST enterotoxin Sth.  
 XX Selection; candidate drug; cell receptor binding; affinity;  
 KW biological receptor; rational drug design; combinatorial drug design;  
 KW receptor antagonist; receptor agonist; ST enterotoxin; beta turn mimetic;  
 KW gamma-turn mimetic; beta sheet mimetic; disulphide bridge mimetic.  
 XX Escherichia coli.  
 OS WO9909416-A2.  
 PN 25-FEB-1999.  
 PD 20-AUG-1998; 98WO-GB002504.  
 XX 20-AUG-1997; 97GB-00017652.  
 PA (NYCO-) NYCOMED IMAGING AS.  
 PA (COCK/) COCKBAIN J.  
 XX Wolfe HR;  
 PI WPI; 1999-181156/15.  
 XX Method of drug selection - and use of an acetamidomethyl-protected  
 PT polymer as a substrate in the solid state synthesis of an oligopeptide.  
 XX Disclosure; Page 1-2; 38pp; English.  
 CC The specification describes a method for selecting a candidate drug  
 CC compound having affinity for biological receptors. The method uses a  
 CC combination of rational and combinatorial drug design techniques. At  
 CC least 1 residue in the original cell receptor binding peptide is modified  
 CC to a non-natural amino acid, preferably a beta turn mimetic, a gamma-turn  
 CC mimetic, a beta sheet mimetic or a disulphide bridge mimetic. The method  
 CC is used for identification of a candidate receptor antagonist or agonist.  
 CC The present peptide is a cell receptor binding peptide, and can thus be  
 CC used as a starting point for identification of candidate drug compounds,  
 CC using the method of the invention  
 XX Sequence 19 AA;

Query Match 78.9%; Score 15; DB 2; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-09;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPAC 15  
 Db 1 NSSNYCCCLCCNPAC 15

RESULT 9  
 AAY29608  
 ID AAY29608 standard; peptide; 19 AA.  
 AC AAY29608;  
 DT 15-OCT-1999 (first entry)  
 DE Escherichia coli heat stable ST enterotoxin STH.  
 KW Heat stable ST enterotoxin; immunoreagent; radiological therapy;  
 KW diagnosis; ST receptor binding moiety; macrocyclic complexing agent;  
 KW tumour; infectious diarrhoeal disease; diarrhoea.  
 XX Escherichia coli.  
 OS WO9939748-A1.  
 PN 12-AUG-1999.  
 PD 08-FEB-1999; 99WO-GB000396.  
 XX

PR 06-FEB-1998; 98US-00020233.  
 XX (NYCO-) NYCOMED IMAGING AS.  
 PA (MATT/) MATTHEWS D P.  
 XX  
 XX Snow RA, Delecki DJ, Shah C, Black C, Wolfe H;  
 PI WPI; 1999-494219/41.  
 DR WPI; 1999-494219/41.  
 XX  
 XX Macrocytic complexing agents containing linked 2,6-pyridiylene nuclei  
 PT as components of targeting immunoreagents binding to ST receptor.  
 XX  
 XX Disclosure; Page 39; 79pp; English.  
 PS  
 XX The present invention describes targeting immunoreagents (TI's)  
 CC comprising a metal ion and a residue of a macrocyclic complexing agent  
 CC (MCA). TI's are of use in diagnostic imaging and therapy of specific  
 CC disease sites in a patient, using either radioactive, magnetic resonance,  
 CC or fluorescent means of detection or use of the metal ion; alternatively,  
 CC a substituent of these types may be introduced, e.g. radioactive iodine,  
 CC to perform the same function. Most notable is the imaging and  
 CC radiological therapy of tumours. In addition, a variety of bacteria,  
 CC including Escherichia coli, Vibrio cholerae, Citrobacter freundii, and  
 CC Yersinia enterocolitica, bind to ST receptors and cause infectious  
 CC diarrhoeal diseases, particularly in pediatrics and in developing  
 CC countries. These types of diarrhoea can also be treated using TI's. TI's  
 CC may specifically used to treat cancers and also be used as an  
 CC antidiarrhoeal agent. TI's are free from the various disadvantages of  
 CC prior art reagents, including rapid destruction and/or excretion,  
 CC instability in storage, and protein degradation. There is no perturbation  
 CC of protein reactive groups at the pyridyl chelating site. AAY29607 to  
 CC AAY29612 represent examples of heat stable ST enterotoxins given in the  
 CC exemplification of the present invention  
 XX  
 XX Sequence 19 AA;  
 SQ

Query Match 78.9%; Score 15; DB 2; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-09;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NSSNYCCELCCNPAC 15  
 DB |||||  
 1 NSSNYCCELCCNPAC 15

RESULT 11  
 AAY02398  
 ID AAY02398 standard; peptide; 19 AA.  
 XX  
 AC AAY02398;  
 XX  
 DT 09-JUL-1999 (first entry)  
 XX  
 DE Heat stable ST enterotoxin Sth.  
 XX  
 KW Selection; candidate drug; cell receptor binding; affinity;  
 KW biological receptor; rational drug design; combinatorial drug design;  
 KW receptor antagonist; receptor agonist; ST enterotoxin; beta turn mimetic;  
 KW gamma-turn mimetic; beta sheet mimetic; disulphide bridge mimetic.  
 XX  
 OS Escherichia coli.  
 XX  
 PN W09909417-A2.  
 XX  
 PD 25-FEB-1999.  
 XX  
 PF 20-AUG-1998; 98WO-GB002510.  
 XX  
 PR 20-AUG-1997; 97GB-00017652.  
 XX  
 XX (NYCO-) NYCOMED IMAGING AS.  
 PA (COCK/) COCKBAIN J.  
 XX  
 XX Wolfe HR;  
 XX  
 XX WPI; 1999-181157/15.  
 XX  
 XX Method of drug selection - using a combination of rational and  
 PT combinatorial drug design techniques.  
 XX  
 XX Disclosure; Page 1-2; 35pp; English.  
 PS  
 XX The specification describes a method for selecting a candidate drug  
 CC

Query Match 78.9%; Score 15; DB 2; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-09;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NSSNYCCELCCNPAC 15  
 DB |||||  
 1 NSSNYCCELCCNPAC 15

RESULT 10  
 AAY06972  
 ID AAY06972 standard; peptide; 19 AA.  
 XX  
 AC AAY06972;  
 XX  
 XX  
 DT 02-JUL-1999 (first entry)  
 XX  
 DE E. coli heat stable ST enterotoxin peptide Sth.  
 XX  
 KW Targeting immunoreagent; metal ion; immunoreactive; terpyridine; tumour;  
 KW complexing agent; diagnostic imaging; radiological treatment; yttrium;  
 KW therapeutic; radiation toxicity; heat stable; ST enterotoxin.  
 XX  
 OS Escherichia coli.  
 XX  
 PN W09921587-A1.  
 XX  
 PD 06-MAY-1999.  
 XX  
 PF 15-OCT-1998; 98WO-GB003102.  
 XX  
 PR 15-OCT-1997; 97US-00951144.  
 XX  
 XX (NYCO-) NYCOMED IMAGING AS.  
 PA (MATT/) MATTHEWS D P.  
 XX  
 XX Wolfe H, Delecki DJ, Yu S;  
 PI

CC compound having affinity for biological receptors. The method uses a  
CC combination of rational and combinatorial drug design techniques. At  
CC least 1 residue in the original cell receptor binding peptide is modified  
CC to a non-natural amino acid, preferably a beta turn mimetic, a gamma-turn  
CC mimetic, a beta sheet mimetic or a disulphide bridge mimetic. The method  
CC is used for identification of a candidate receptor antagonist or agonist.  
CC The present peptide is a cell receptor binding peptide, and can thus be  
CC used as a starting point for identification of candidate drug compounds,  
CC using the method of the invention  
XX  
SQ Sequence 19 AA;

Query Match 78.9%; Score 15; DB 2; Length 19;  
Best Local Similarity 100.0%; Pred. No. 2.2e-09;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NSSNYCCCLCCNPAC 15  
DB 1 NSSNYCCCLCCNPAC 15  
|||||

RESULT 12  
AAMS1879  
ID AAMS1879 standard; peptide; 19 AA.

AC AAMS1879;  
XX  
DT 01-FEB-2002 (first entry)  
XX  
DE Human thermostable enterotoxin Sth peptide fragment #3.  
KW Human; thermostable enterotoxin; Sth; metastatic colorectal cancer;  
KW guanyl cyclase-C; GC-C; Sth.

XX Homo sapiens.  
XX PR2805994-A1.  
XX  
PD 14-SEP-2001.

PF 10-MAR-2000; 2000PR-00003141.  
XX  
PR 10-MAR-2000; 2000PR-00003141.  
XX  
PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.

XX Der Vartanian M, Batisson I;  
XX WPI; 2001-640835/74.

XX New compound for detecting and treating metastatic colorectal cancer  
PT comprises a conjugate of an Sth peptide and an immunogenic protein which  
PT binds to the guanyl cyclase-c receptor.

PS Disclosure; Page 44; 126pp; French.

XX The present invention relates to a conjugate which comprises an E. coli  
CC thermostable enterotoxin (Sth) peptide and an active molecule where the  
CC Sth peptide has a conformation such that it is capable of binding to the  
CC guanyl cyclase-C (GC-C) receptor. This can be used in the specific  
CC diagnosis and treatment of metastatic colorectal cancer. The present  
CC sequence is a fragment of the human thermostable enterotoxin (Sth)  
XX protein

SQ Sequence 19 AA;

Query Match 78.9%; Score 15; DB 4; Length 19;  
Best Local Similarity 100.0%; Pred. No. 2.2e-09;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPAC 15  
DB 1 NSSNYCCCLCCNPAC 15  
|||||

RESULT 13  
AAO16204  
ID AAO16204 standard; peptide; 19 AA.

XX AAO16204;  
XX  
DT 28-MAR-2003 (first entry)

XX E coli small heat stable enterotoxin tricyclo peptide, SEQ ID No 23.

XX Guanylate cyclase receptor agonist; apoptosis induction; cancer; polyps;  
KW inflammation; asthma; nephritis; hepatitis; bronchitis; cystic fibrosis;  
KW small heat stable enterotoxin tricyclo; inflammatory bowel disease;  
KW pancreatitis; ulcerative colitis; Crohn's disease; Kaposi's sarcoma.

XX Escherichia coli.

XX Key Location/Qualifiers  
FH Disulfide-bond 6..10  
FT Disulfide-bond 7..15  
FT Disulfide-bond 11..18

XX WO200278683-A1.

XX 10-OCT-2002.

XX 28-MAR-2002; 2002WO-US009551.

XX 29-MAR-2001; 2001US-0279437P.

XX 27-JUN-2001; 2001US-0300850P.

XX 10-JUL-2001; 2001US-0303808P.

XX 25-JUL-2001; 2001US-0307358P.

XX 17-JAN-2002; 2002US-0348646P.

XX (SYNE-) SYNERGY PHARM INC.

XX Shailubhai K, Nikiforovich G, Jacob GS;

XX WPI; 2003-148251/14.  
XX Novel guanylate cyclase receptor agonist peptide useful for preventing or  
PT treating primary or metastatic cancer and polyps in a patient, and for  
PT inducing apoptosis in the cells of a subject.

XX Disclosure; Page 10; 47pp; English.

XX The invention comprises guanylate cyclase receptor agonist peptides that  
CC are useful for inducing apoptosis in the cells of a subject. The peptides  
CC of the invention may be used to treat: cancer; polyps; inflammation;  
CC asthma; nephritis; hepatitis; pancreatitis; bronchitis; cystic fibrosis;  
CC inflammatory bowel disease; ulcerative colitis; Crohn's disease; and  
CC Kaposi's sarcoma. The present amino acid sequence represents an  
CC Escherichia coli small heat stable enterotoxin tricyclo peptide which was  
CC used in the invention

XX Sequence 19 AA;

Query Match 78.9%; Score 15; DB 6; Length 19;  
Best Local Similarity 100.0%; Pred. No. 2.2e-09;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPAC 15  
DB 1 NSSNYCCCLCCNPAC 15  
|||||

RESULT 14  
ABG74825  
ID ABG74825 standard; peptide; 19 AA.

XX

AC ABG74825;  
XX 12-JUN-2003 (first entry)  
XX E. coli heat stable enterotoxin derived peptide SEQ ID 7.  
XX  
XX Apical membrane; mucosal epithelial cell; respiratory tract;  
KW guanylate cyclase C; G protein-coupled receptor; guanosine triphosphate;  
KW cyclic guanosine monophosphate; cGMP; chloride ion secretion; inhalation;  
KW membrane-associated type II protein kinase; mucus fluidisation;  
KW cystic fibrosis transmembrane conductance regulator; breathing disorder;  
KW mucus secretion; antiasthmatic; antiinflammatory; bronchial asthma;  
KW chronic bronchitis; cystic fibrosis; enterotoxin; heat stable.  
XX Escherichia coli.  
OS  
XX WO200298912-A2.  
XX  
XX 12-DEC-2002.  
XX  
XX 05-JUN-2002; 2002WO-DE002040.  
XX  
XX 05-JUN-2001; 2001DE-01027119.  
XX  
XX (CETI/) CETIN Y.  
XX (SAVA/) SAVAS Y.  
XX  
XX Cetin Y, Savas Y;  
XX  
XX WPI; 2003-156842/15.  
XX  
XX Composition useful for treating respiratory disease, comprises a peptide  
PT that activates guanylate cyclase C, and is delivered to the apical  
PT membrane through the respiratory tract.  
XX  
XX Claim 3; Page 4; 23pp; German.  
XX  
XX This invention describes a novel medicament in a formulation that is  
CC delivered to the apical membrane of mucosal epithelial cells through the  
CC respiratory tract. The medicament contains at least one peptide that  
CC activates guanylate cyclase C (GCC). GCC is a G protein-coupled receptor  
CC that catalyses conversion of guanosine triphosphate to cyclic guanosine  
CC monophosphate (cGMP) and is present on the apical (air) side of  
CC respiratory epithelial cells but not on the basolateral (blood) side.  
CC cGMP activates membrane-associated type II protein kinase which in turn  
CC activates the regulatory domain of the cystic fibrosis transmembrane  
CC conductance regulator, resulting in secretion of chloride ions and water  
CC from the cells, causing fluidisation of the mucus. The products of the  
CC invention are used to make an inhalation device containing the medicament  
CC for diagnosing diseases that are accompanied by breathing disorders or  
CC disorders of mucus secretion in the respiratory tract, by detecting at  
CC least one GCC activator. The products of the invention have antiasthmatic  
CC and antiinflammatory activity. The method is useful for diagnosing and  
CC treating diseases accompanied by breathing disorders or disorders of  
CC mucus secretion in the respiratory tract particularly bronchial asthma,  
CC chronic bronchitis and cystic fibrosis. The product of the invention  
CC improves fluidity and evacuation of bronchial mucus and acts locally  
CC (since the medicament does not enter the bloodstream), so systemic side  
CC effects are minimised. Only very small doses of the medicament are  
CC required. This sequence represents a heat stable E. coli enterotoxin  
CC derived peptide use in an assay to determine guanylate cyclase C  
CC activation described in the disclosure of the invention  
XX  
SQ Sequence 19 AA;  
Query Match 78.9%; Score 15; DB 6; Length 19;  
Best Local Similarity 100.0%; Pred. No. 2.2e-09;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NSSNYCCCLCCNPAC 15  
Db 1 NSSNYCCCLCCNPAC 15  
Search completed: March 26, 2005, 17:20:58  
Job time : 85.4821 secs

RESULT 15  
ADCI14120  
ID ADCI14120 standard; peptide; 19 AA.  
XX  
XX ADCI14120;  
AC  
XX 18-DEC-2003 (first entry)  
DT  
XX Heat stable enterotoxin, Sth.  
DE  
XX Breast cancer; prostate cancer; pancreatic cancer; melanoma;  
KW heat stable enterotoxin; ST motif; tail region; receptor binding region;  
KW cytotostatic.  
KW  
XX Escherichia coli.  
OS  
XX  
XX Key Location/Qualifiers  
FH Disulfide-bond 6..11  
FT Disulfide-bond 7..15  
FT Disulfide-bond 10..18.  
XX  
XX WO2003072125-A1.  
PN  
XX 04-SEP-2003.  
PD  
XX 21-FEB-2003; 2003WO-US005343.  
PF  
XX 22-FEB-2002; 2002US-0359204P.  
PR  
XX (UMOR ) UNIV MISSOURI.  
PA  
XX Sieckman G, Gali H, Volkert W, Forte L, Hoffman T;  
PI  
XX WPI; 2003-731571/69.  
DR  
XX  
XX Method of treating or preventing breast, prostate, pancreatic cancer or  
PT melanoma, comprises targeting peptide agent complex to breast cancer  
PT cell, prostate cancer cell, pancreatic cancer cell or melanoma cancer  
PT cell.  
XX  
XX Disclosure; Fig 1; 4lpp; English.  
XX  
XX The present invention relates to a method for targeting an agent to a  
CC breast cancer cell, prostate cancer cell, a pancreatic cancer cell or  
CC melanoma cancer cell. The method involves contacting the cancer cell with  
CC a peptide agent complex wherein the peptide comprises an E. coli heat  
CC stable enterotoxin (ST) motif that binds to the cancer cell. The ST motif  
CC may comprise a tail region (ADCI14118) and an receptor binding region  
CC (ADCI14119). The method is used for targeting an agent to breast,  
CC prostate, pancreatic or melanoma cancer cell; for rendering, treating and  
CC preventing an unresectable breast, prostate, pancreatic or melanoma  
CC tumour and for identifying tumour binding peptide. The present sequence  
CC is heat stable enterotoxin (Sth) produced by human strain of Escherichia  
CC coli bacteria.  
XX  
SQ Sequence 19 AA;  
Query Match 78.9%; Score 15; DB 7; Length 19;  
Best Local Similarity 100.0%; Pred. No. 2.2e-09;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NSSNYCCCLCCNPAC 15  
Db 1 NSSNYCCCLCCNPAC 15

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2005, 17:12:01 ; Search time 25.7857 Seconds  
(without alignments)  
55.005 Million cell updates/sec

Title: US-10-775-481A-5

Perfect score: 19

Sequence: 1 NSSNYCCELCCNPACNGCY 19

Scoring table:  Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 19    | 100.0       | 19     | 1  | US-08-141-892A-5  |
| 2          | 19    | 100.0       | 19     | 2  | US-08-583-447A-5  |
| 3          | 19    | 100.0       | 19     | 2  | US-08-467-920-5   |
| 4          | 19    | 100.0       | 19     | 3  | US-08-635-930-5   |
| 5          | 19    | 100.0       | 19     | 3  | US-09-193-997-5   |
| 6          | 19    | 100.0       | 19     | 3  | US-09-138-237A-5  |
| 7          | 15    | 78.9        | 18     | 1  | US-08-141-892A-27 |
| 8          | 15    | 78.9        | 18     | 2  | US-08-583-447A-27 |
| 9          | 15    | 78.9        | 18     | 2  | US-08-467-920-27  |
| 10         | 15    | 78.9        | 18     | 3  | US-08-635-930-27  |
| 11         | 15    | 78.9        | 18     | 3  | US-09-193-997-27  |
| 12         | 15    | 78.9        | 18     | 3  | US-09-138-237A-27 |
| 13         | 14    | 73.7        | 17     | 1  | US-08-141-892A-28 |
| 14         | 14    | 73.7        | 17     | 2  | US-08-583-447A-28 |
| 15         | 14    | 73.7        | 17     | 2  | US-08-467-920-28  |
| 16         | 14    | 73.7        | 17     | 3  | US-08-635-930-28  |
| 17         | 14    | 73.7        | 17     | 3  | US-09-193-997-28  |
| 18         | 14    | 73.7        | 17     | 3  | US-09-138-237A-28 |
| 19         | 14    | 73.7        | 18     | 1  | US-08-141-892A-33 |
| 20         | 14    | 73.7        | 18     | 2  | US-08-583-447A-33 |
| 21         | 14    | 73.7        | 18     | 2  | US-08-467-920-33  |
| 22         | 14    | 73.7        | 18     | 3  | US-08-635-930-33  |
| 23         | 14    | 73.7        | 18     | 3  | US-09-193-997-33  |
| 24         | 14    | 73.7        | 18     | 3  | US-09-138-237A-33 |
| 25         | 13    | 68.4        | 16     | 1  | US-08-141-892A-29 |
| 26         | 13    | 68.4        | 16     | 2  | US-08-583-447A-29 |
| 27         | 13    | 68.4        | 16     | 2  | US-08-467-920-29  |

|    |    |      |    |   |                   |                   |
|----|----|------|----|---|-------------------|-------------------|
| 28 | 13 | 68.4 | 16 | 3 | US-08-635-930-29  | Sequence 29, Appl |
| 29 | 13 | 68.4 | 16 | 3 | US-09-193-997-29  | Sequence 29, Appl |
| 30 | 13 | 68.4 | 16 | 3 | US-09-138-237A-29 | Sequence 29, Appl |
| 31 | 13 | 68.4 | 17 | 1 | US-08-141-892A-34 | Sequence 34, Appl |
| 32 | 13 | 68.4 | 17 | 2 | US-08-583-447A-34 | Sequence 34, Appl |
| 33 | 13 | 68.4 | 17 | 2 | US-08-467-920-34  | Sequence 34, Appl |
| 34 | 13 | 68.4 | 17 | 3 | US-08-635-930-34  | Sequence 34, Appl |
| 35 | 13 | 68.4 | 17 | 3 | US-09-193-997-34  | Sequence 34, Appl |
| 36 | 13 | 68.4 | 17 | 3 | US-09-138-237A-34 | Sequence 34, Appl |
| 37 | 12 | 63.2 | 15 | 1 | US-08-141-892A-30 | Sequence 30, Appl |
| 38 | 12 | 63.2 | 15 | 2 | US-08-583-447A-30 | Sequence 30, Appl |
| 39 | 12 | 63.2 | 15 | 2 | US-08-467-920-30  | Sequence 30, Appl |
| 40 | 12 | 63.2 | 15 | 3 | US-08-635-930-30  | Sequence 30, Appl |
| 41 | 12 | 63.2 | 15 | 3 | US-09-193-997-30  | Sequence 30, Appl |
| 42 | 12 | 63.2 | 15 | 3 | US-09-138-237A-30 | Sequence 30, Appl |
| 43 | 12 | 63.2 | 16 | 1 | US-08-141-892A-35 | Sequence 35, Appl |
| 44 | 12 | 63.2 | 16 | 2 | US-08-583-447A-35 | Sequence 35, Appl |
| 45 | 12 | 63.2 | 16 | 2 | US-08-467-920-35  | Sequence 35, Appl |

ALIGNMENTS

RESULT 1  
US-08-141-892A-5  
; Sequence 5, Application US/08141892A  
; Patent No. 5518888  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: ST Receptor Binding Compounds and Methods  
; TITLE OF INVENTION: of Using the Same  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 5518888Bris  
; STREET: One Liberty Place - 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: U.S.A.  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch disk, 720 Kb  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/141,892A  
; FILING DATE: 26-OCT-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: TJU-0903  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-141-892A-5

Query Match 100.0%; Score 19; DB 1; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.3e-13;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSSNYCCELCCNPACNGCY 19  
Db 1 NSSNYCCELCCNPACNGCY 19

RESULT 2  
US-08-583-447A-5  
; Sequence 5, Application US/08583447A  
; Patent No. 5879656  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: ST Receptor Binding Compounds and  
; TITLE OF INVENTION: Methods of Using the Same  
; NUMBER OF SEQUENCES: 56  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5879656ris  
; STREET: One Liberty Place, 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: WordPerfect 6.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/583,447A  
; FILING DATE: 05-JAN-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/141,892  
; FILING DATE: 26-OCT-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DeLuca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: TJU-1702  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-583-447A-5  
Query Match 100.0%; Score 19; DB 2; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.3e-13;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NSSNYCCCLCCNPACNGCY 19  
Db 1 NSSNYCCCLCCNPACNGCY 19

RESULT 3  
US-08-467-920-5  
; Sequence 5, Application US/08467920  
; Patent No. 5962220  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: Compositions That Specifically  
; TITLE OF INVENTION: Bind To Colorectal Cancer Cells  
; TITLE OF INVENTION: And Methods Of Using The Same  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; STREET: One Liberty Place, 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19103

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/467,920  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/141,892  
; FILING DATE: 26-OCT-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DeLuca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: TJU-1589  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-467-920-5  
Query Match 100.0%; Score 19; DB 2; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.3e-13;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NSSNYCCCLCCNPACNGCY 19  
Db 1 NSSNYCCCLCCNPACNGCY 19

RESULT 4  
US-08-635-930-5  
; Sequence 5, Application US/08635930  
; Patent No. 6060037  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: Compositions That Specifically Bind To  
; TITLE OF INVENTION: Colorectal Cancer Cells And Methods Of Using  
; TITLE OF INVENTION: The Same  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6060037ris  
; STREET: One Liberty Place, 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: WINDOWS 3.1  
; SOFTWARE: WordPerfect 6.0/6.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/635,930  
; FILING DATE: 26-APR-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/141,892  
; FILING DATE: 26-OCT-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/305,056  
; FILING DATE: 13-SEP-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DeLuca, Mark



REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-1360  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-635-930-5

Query Match 100.0%; Score 19; DB 3; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.3e-13;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSSNYCCCLCCNPACNGCY 19  
Db 1 NSSNYCCCLCCNPACNGCY 19

## RESULT 5

US-09-193-997-5  
Sequence 5, Application US/09193997

PATENT NO. 6087109  
GENERAL INFORMATION:  
APPLICANT: Waldman, Scott A.  
TITLE OF INVENTION: Compositions That Specifically  
Bind To Colorectal Cancer Cells  
TITLE OF INVENTION: Bind To Colorectal Cancer Cells  
TITLE OF INVENTION: And Methods Of Using The Same  
NUMBER OF SEQUENCES: 54

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
ADDRESSEE: No. 6087109ris  
STREET: One Liberty Place, 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/193,997

FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/467,920

FILING DATE:

CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:

NAME: DeLuca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-1589  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-193-997-5

Query Match 100.0%; Score 19; DB 3; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.3e-13;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSSNYCCCLCCNPACNGCY 19

Db 1 NSSNYCCCLCCNPACNGCY 19

## RESULT 6

US-09-138-237A-5  
Sequence 5, Application US/09138237A

PATENT NO. 6268159  
GENERAL INFORMATION:  
APPLICANT: Waldman, Scott A.  
TITLE OF INVENTION: ST Receptor Binding Compounds and Methods  
TITLE OF INVENTION: Of Using the Same  
NUMBER OF SEQUENCES: 54

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 6268159ris  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19103

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch disk, 720 Kb  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/138,237A

FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/141,892

FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: DeLuca, Mark

REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-0903

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:

LENGTH: 19 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein

US-09-138-237A-5

Query Match 100.0%; Score 19; DB 3; Length 19;

Best Local Similarity 100.0%; Pred. No. 1.3e-13;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSSNYCCCLCCNPACNGCY 19

Db 1 NSSNYCCCLCCNPACNGCY 19

## RESULT 7

US-08-141-892A-27

Sequence 27, Application US/08141892A  
Patent No. 5518888

GENERAL INFORMATION:

APPLICANT: Waldman, Scott A.

TITLE OF INVENTION: ST Receptor Binding Compounds and Methods  
TITLE OF INVENTION: Of Using the Same

NUMBER OF SEQUENCES: 54

CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 5518888ris  
STREET: One Liberty Place - 46th Floor

CITY: Philadelphia  
STATE: Pennsylvania

COUNTRY: U.S.A.  
ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch disk, 720 Kb  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/141,892A  
FILING DATE: 26-OCT-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Deluca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-0903  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-141-892A-27

Query Match 78.9%; Score 15; DB 1; Length 18;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPAC 15  
DB 1 NSSNYCCCLCCNPAC 15

## RESULT 8

US-08-583-447A-27  
Sequence 27, Application US/08583447A  
Patent No. 5879656  
GENERAL INFORMATION:

APPLICANT: Waldman, Scott A.  
TITLE OF INVENTION: St Receptor Binding Compounds and  
METHODS OF USING THE SAME  
NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5879656ris  
STREET: One Liberty Place, 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: WordPerfect 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/583,447A  
FILING DATE: 05-JAN-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/141,892  
FILING DATE: 26-OCT-1993  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Deluca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-1702  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:

LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-583-447A-27

Query Match 78.9%; Score 15; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPAC 15  
DB 1 NSSNYCCCLCCNPAC 15

## RESULT 9

US-08-467-920-27  
Sequence 27, Application US/08467920  
Patent No. 5962220  
GENERAL INFORMATION:

APPLICANT: Waldman, Scott A.  
TITLE OF INVENTION: Compositions That Specifically  
Bind To Colorectal Cancer Cells  
METHODS OF USING THE SAME  
NUMBER OF SEQUENCES: 54  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
STREET: One Liberty Place, 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,920  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/141,892  
FILING DATE: 26-OCT-1993  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Deluca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-1589  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-467-920-27

Query Match 78.9%; Score 15; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPAC 15  
DB 1 NSSNYCCCLCCNPAC 15

## RESULT 10

US-08-635-930-27  
Sequence 27, Application US/08635930

Patent No. 6060037  
GENERAL INFORMATION:  
APPLICANT: Waldman, Scott A.  
TITLE OF INVENTION: Compositions That Specifically Bind To  
OPERATING SYSTEM: Colorectal Cancer Cells And Methods Of Using  
TITLE OF INVENTION: The Same  
NUMBER OF SEQUENCES: 54  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6060037ris  
STREET: One Liberty Place, 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: WINDOWS 3.1  
SOFTWARE: WordPerfect 6.0/6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/635,930  
FILING DATE: 26-APR-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/141,892  
FILING DATE: 26-OCT-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/305,056  
FILING DATE: 13-SEP-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: DeLuca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-1360  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-635-930-27

Query March 78.9%; Score 15; DB 3; Length 18;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSSNYCCCLCCNPAC 15  
Db 1 NSSNYCCCLCCNPAC 15

RESULT 11  
US-09-193-997-27  
Sequence 27, Application US/09193997  
Patent No. 6087109  
GENERAL INFORMATION:  
APPLICANT: Waldman, Scott A.  
TITLE OF INVENTION: Compositions That Specifically  
OPERATING SYSTEM: Bind To Colorectal Cancer Cells  
TITLE OF INVENTION: And Methods Of Using The Same  
NUMBER OF SEQUENCES: 54  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
STREET: One Liberty Place, 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/193,997  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/467,920  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: DeLuca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-1589  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-193-997-27

Query March 78.9%; Score 15; DB 3; Length 18;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSSNYCCCLCCNPAC 15  
Db 1 NSSNYCCCLCCNPAC 15

RESULT 12  
US-09-138-237A-27  
Sequence 27, Application US/09138237A  
Patent No. 6268159  
GENERAL INFORMATION:  
APPLICANT: Waldman, Scott A.  
TITLE OF INVENTION: St Receptor Binding Compounds and Methods  
TITLE OF INVENTION: of Using the Same  
NUMBER OF SEQUENCES: 54  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 6268159ris  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch disk, 720 Kb  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/138,237A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/141,892  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: DeLuca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-0903  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-138-237A-27

Query Match 78.9%; Score 15; DB 3; Length 18;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPAC 15  
Db 1 NSSNYCCCLCCNPAC 15

## RESULT 13

US-08-141-892A-28  
Sequence 28, Application US/08141892A  
Patent No. 5518888  
GENERAL INFORMATION:  
APPLICANT: Waldman, Scott A.  
TITLE OF INVENTION: ST Receptor Binding Compounds and Methods  
TITLE OF INVENTION: Of Using the Same  
NUMBER OF SEQUENCES: 54  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 5518888ris  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19103

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch disk, 720 Kb  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/141,892A  
FILING DATE: 26-OCT-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Deluca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-0903  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-141-892A-28

Query Match 73.7%; Score 14; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1.9e-08;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSNYCCCLCCNPAC 15  
Db 1 SSNYCCCLCCNPAC 14

## RESULT 14

US-08-583-447A-28  
Sequence 28, Application US/08583447A  
Patent No. 5879656  
GENERAL INFORMATION:

APPLICANT: Waldman, Scott A.  
TITLE OF INVENTION: ST Receptor Binding Compounds and  
TYPE: Methods of Using the Same  
NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5879656ris  
STREET: One Liberty Place, 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: WordPerfect 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/583,447A  
FILING DATE: 05-JAN-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/141,892  
FILING DATE: 26-OCT-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Deluca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-1702  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-583-447A-28

Query Match 73.7%; Score 14; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1.9e-08;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSNYCCCLCCNPAC 15  
Db 1 SSNYCCCLCCNPAC 14

## RESULT 15

US-08-467-920-28  
Sequence 28, Application US/08467920  
Patent No. 5962220  
GENERAL INFORMATION:  
APPLICANT: Waldman, Scott A.  
TITLE OF INVENTION: Compositions That Specifically  
Bind To Colorectal Cancer Cells  
TITLE OF INVENTION: And Methods Of Using The Same  
NUMBER OF SEQUENCES: 54  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
ADDRESS: No. 5962220ris  
STREET: One Liberty Place, 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,920

```

; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/141,892
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1589
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-467-920-28

```

```

Query Match      73.7%; Score 14; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      2 SSNYCCCLCCNPAC 15
        |||||
Db      1 SSNYCCCLCCNPAC 14

```

```

Search completed: March 26, 2005, 17:27:27
Job time : 26.7857 secs

```

***This Page Blank (uspto)***

Pending Nucleic Acid and Pending Amino Acid database searches generate two sets of results each. The Pending databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches. Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions .rnpm and .rnpn. Searches run against the Amino Acid Pending database produce two sets of results, with the extensions .rapm and .rapn.

***Because they contain data that is confidential, the results of Pending database searches should not be left in the case.***

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2005, 17:25:02 ; Search time 61.4107 Seconds  
(without alignments)  
102.440 Million cell updates/sec

Title: US-10-775-481A-5

Perfect score: 19

Sequence: 1 NSSNYCCELCNACNGCY 19

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 1407402 seqs, 331100923 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1407402

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description      |
|------------|-------|-------------|--------|----|------------------|
| 1          | 19    | 100.0       | 19     | 15 | US-10-621-684-5  |
| 2          | 15    | 78.9        | 18     | 15 | US-10-621-684-27 |
| 3          | 15    | 78.9        | 19     | 14 | US-10-107-814-23 |
| 4          | 15    | 78.9        | 19     | 15 | US-10-371-966-1  |
| 5          | 15    | 78.9        | 19     | 15 | US-10-371-966-2  |
| 6          | 15    | 78.9        | 19     | 17 | US-10-796-719-1  |
| 7          | 15    | 78.9        | 19     | 17 | US-10-796-719-26 |
| 8          | 15    | 78.9        | 19     | 17 | US-10-796-719-27 |
| 9          | 15    | 78.9        | 19     | 17 | US-10-479-606-7  |
| 10         | 15    | 78.9        | 21     | 17 | US-10-796-719-39 |
| 11         | 15    | 78.9        | 21     | 17 | US-10-796-719-40 |
| 12         | 15    | 78.9        | 72     | 17 | US-10-796-719-21 |
| 13         | 14    | 73.7        | 17     | 15 | US-10-621-684-28 |

|    |      |    |    |                  |                   |
|----|------|----|----|------------------|-------------------|
| 14 | 73.7 | 18 | 15 | US-10-621-684-33 | Sequence 33, Appl |
| 15 | 68.4 | 16 | 15 | US-10-621-684-29 | Sequence 29, Appl |
| 16 | 68.4 | 17 | 15 | US-10-621-684-34 | Sequence 34, Appl |
| 17 | 63.2 | 15 | 15 | US-10-621-684-30 | Sequence 30, Appl |
| 18 | 63.2 | 15 | 15 | US-10-621-684-35 | Sequence 35, Appl |
| 19 | 57.9 | 14 | 15 | US-10-621-684-11 | Sequence 11, Appl |
| 20 | 57.9 | 14 | 15 | US-10-621-684-31 | Sequence 31, Appl |
| 21 | 57.9 | 14 | 15 | US-10-621-684-43 | Sequence 43, Appl |
| 22 | 57.9 | 15 | 15 | US-10-621-684-10 | Sequence 10, Appl |
| 23 | 57.9 | 15 | 15 | US-10-621-684-16 | Sequence 16, Appl |
| 24 | 57.9 | 15 | 15 | US-10-621-684-36 | Sequence 36, Appl |
| 25 | 57.9 | 15 | 15 | US-10-621-684-42 | Sequence 42, Appl |
| 26 | 57.9 | 16 | 15 | US-10-621-684-9  | Sequence 9, Appl  |
| 27 | 57.9 | 16 | 15 | US-10-621-684-15 | Sequence 15, Appl |
| 28 | 57.9 | 17 | 15 | US-10-621-684-8  | Sequence 8, Appl  |
| 29 | 57.9 | 17 | 15 | US-10-621-684-14 | Sequence 14, Appl |
| 30 | 57.9 | 18 | 15 | US-10-621-684-41 | Sequence 41, Appl |
| 31 | 57.9 | 18 | 15 | US-10-621-684-7  | Sequence 7, Appl  |
| 32 | 57.9 | 18 | 15 | US-10-621-684-13 | Sequence 13, Appl |
| 33 | 57.9 | 18 | 15 | US-10-621-684-38 | Sequence 38, Appl |
| 34 | 57.9 | 18 | 17 | US-10-796-719-2  | Sequence 2, Appl  |
| 35 | 57.9 | 18 | 17 | US-10-796-719-4  | Sequence 4, Appl  |
| 36 | 57.9 | 18 | 17 | US-10-796-719-11 | Sequence 11, Appl |
| 37 | 57.9 | 19 | 15 | US-10-621-684-2  | Sequence 2, Appl  |
| 38 | 57.9 | 19 | 15 | US-10-621-684-49 | Sequence 49, Appl |
| 39 | 57.9 | 72 | 17 | US-10-796-719-20 | Sequence 20, Appl |
| 40 | 52.6 | 13 | 15 | US-10-621-684-12 | Sequence 12, Appl |
| 41 | 52.6 | 13 | 15 | US-10-621-684-32 | Sequence 32, Appl |
| 42 | 52.6 | 13 | 15 | US-10-621-684-45 | Sequence 45, Appl |
| 43 | 52.6 | 14 | 15 | US-10-621-684-17 | Sequence 17, Appl |
| 44 | 52.6 | 14 | 15 | US-10-621-684-37 | Sequence 37, Appl |
| 45 | 52.6 | 14 | 15 | US-10-621-684-44 | Sequence 44, Appl |

#### ALIGNMENTS

#### RESULT 1

US-10-621-684-5

; Sequence 5, Application US/10621684

; Publication No. US20040029182A1

; GENERAL INFORMATION:

; APPLICANT: Waldman, Scott A.

; TITLE OF INVENTION: ST Receptor Binding Compounds and

; Methods of Using the Same

; NUMBER OF SEQUENCES: 56

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. US20040029182A1ris

; STREET: One Liberty Place, 46th Floor

; CITY: Philadelphia

; STATE: Pennsylvania

; COUNTRY: USA

; ZIP: 19103

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: Windows

; SOFTWARE: Wordperfect 6.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/621,684

; FILING DATE: 17-Jul-2003

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/583,447A

; FILING DATE: 05-JAN-1996

; APPLICATION NUMBER: US 08/141,892

; FILING DATE: 26-OCT-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: DeLuca, Mark

; REGISTRATION NUMBER: 33,229

; REFERENCE/DOCKET NUMBER: TLU-1702

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 215-568-3100

TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-10-621-684-5

Query Match 100.0%; Score 19; DB 15; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.2e-13;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPACNGCY 19  
DB 1 NSSNYCCCLCCNPACNGCY 19

RESULT 2  
US-10-621-684-27  
; Sequence 27, Application US/10621684  
; Publication No. US20040029182A1  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: ST Receptor Binding Compounds and  
; METHODS OF USING THE SAME  
; NUMBER OF SEQUENCES: 56  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. US20040029182A1ris  
; STREET: One Liberty Place, 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19103

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: WordPerfect 6.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/621.684  
; FILING DATE: 17-Jul-2003  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/583.447A  
; FILING DATE: 05-JAN-1996  
; APPLICATION NUMBER: US 08/141,892  
; FILING DATE: 26-OCT-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DeLuca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: TJU-1702  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:  
US-10-621-684-27

Query Match 78.9%; Score 15; DB 15; Length 18;  
Best Local Similarity 100.0%; Pred. No. 2.3e-09;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPAC 15  
DB 1 NSSNYCCCLCCNPAC 15

RESULT 3  
US-10-107-814-23  
; Sequence 23, Application US/10107814  
; Publication No. US20030073628A1  
; GENERAL INFORMATION:  
; APPLICANT: SHAILUBHAI, KUNWAR  
; APPLICANT: NIKIFOROVICH, GREGORY  
; APPLICANT: JACOB, GARY S.  
; TITLE OF INVENTION: GUANYLATE CYCLASE RECEPTOR AGONISTS FOR THE TREATMENT  
; OF TISSUE INFLAMMATION AND CARCINOGENESIS  
; FILE REFERENCE: 81361/284943/MAS  
; CURRENT APPLICATION NUMBER: US/10/107.814  
; CURRENT FILING DATE: 2002-03-28  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 23  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
; FEATURE:  
; NAME/KEY: DISULFID  
; LOCATION: (6)..(10)  
; NAME/KEY: DISULFID  
; LOCATION: (7)..(15)  
; NAME/KEY: DISULFID  
; LOCATION: (11)..(18)  
US-10-107-814-23

Query Match 78.9%; Score 15; DB 14; Length 19;  
Best Local Similarity 100.0%; Pred. No. 2.4e-09;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPAC 15  
DB 1 NSSNYCCCLCCNPAC 15

RESULT 4  
US-10-371-966-1  
; Sequence 1, Application US/10371966  
; Publication No. US20030232013A1  
; GENERAL INFORMATION:  
; APPLICANT: SIECKMAN, GARY  
; APPLICANT: VOLKERT, WYNN  
; APPLICANT: FORTE, LEONARD  
; APPLICANT: HOFFMAN, TIMOTHY  
; APPLICANT: GALLI, HARIPRASAD  
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TARGETING OF CANCERS CELLS  
; WITH TUMOR HOMING PEPTIDES  
; FILE REFERENCE: UVMO:013US  
; CURRENT APPLICATION NUMBER: US/10/371.966  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/359,204  
; PRIOR FILING DATE: 2002-02-22  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Peptide  
US-10-371-966-1

Query Match 78.9%; Score 15; DB 15; Length 19;  
Best Local Similarity 100.0%; Pred. No. 2.4e-09;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPAC 15  
DB 1 NSSNYCCCLCCNPAC 15

```
RESULT 5
US-10-371-966-2
; Sequence 2, Application US/10371966
; Publication No. US20030232013A1
; GENERAL INFORMATION:
; APPLICANT: SIECKMAN, GARY
; APPLICANT: VOLKERT, WYNN
; APPLICANT: FORTE, LEONARD
; APPLICANT: HOFFMAN, TIMOTHY
; APPLICANT: GALI, HARIPRASAD
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TARGETING OF CANCERS CELLS
; TITLE OF INVENTION: WITH TUMOR HOMING PEPTIDES
; FILE REFERENCE: UWO:01305
; CURRENT APPLICATION NUMBER: US/10/371.966
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/359,204
; PRIOR FILING DATE: 2002-02-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; TYPE: PRT
; LENGTH: 19
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-371-966-2

Query Match      78.9%; Score 15; DB 15; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 NSSNYCCCLCCNPAC 15
Db      1 NSSNYCCCLCCNPAC 15

RESULT 6
US-10-796-719-1
; Sequence 1, Application US/10796719
; Publication No. US20050020811A1
; GENERAL INFORMATION:
; APPLICANT: Currie, Mark G.
; APPLICANT: Mahajan-Miklos, Shalina
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TITLE OF INVENTION: TREATMENT OF GASTROINTESTINAL DISORDERS
; FILE REFERENCE: 14184-043001
; CURRENT APPLICATION NUMBER: US/10/796.719
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US 10/766,735
; PRIOR FILING DATE: 2004-01-28
; PRIOR APPLICATION NUMBER: US 60/443,098
; PRIOR FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 60/471,288
; PRIOR FILING DATE: 2003-05-15
; PRIOR APPLICATION NUMBER: US 60/519,460
; PRIOR FILING DATE: 2003-11-12
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-796-719-1

Query Match      78.9%; Score 15; DB 17; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 NSSNYCCCLCCNPAC 15
Db      1 NSSNYCCCLCCNPAC 15

RESULT 7
US-10-796-719-26
; Sequence 26, Application US/10796719
; Publication No. US20050020811A1
; GENERAL INFORMATION:
; APPLICANT: Currie, Mark G.
; APPLICANT: Mahajan-Miklos, Shalina
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TITLE OF INVENTION: TREATMENT OF GASTROINTESTINAL DISORDERS
; FILE REFERENCE: 14184-043001
; CURRENT APPLICATION NUMBER: US/10/796.719
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US 10/766,735
; PRIOR FILING DATE: 2004-01-28
; PRIOR APPLICATION NUMBER: US 60/443,098
; PRIOR FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 60/471,288
; PRIOR FILING DATE: 2003-05-15
; PRIOR APPLICATION NUMBER: US 60/519,460
; PRIOR FILING DATE: 2003-11-12
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated peptide
US-10-796-719-26

Query Match      78.9%; Score 15; DB 17; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 NSSNYCCCLCCNPAC 15
Db      1 NSSNYCCCLCCNPAC 15

RESULT 8
US-10-796-719-27
; Sequence 27, Application US/10796719
; Publication No. US20050020811A1
; GENERAL INFORMATION:
; APPLICANT: Currie, Mark G.
; APPLICANT: Mahajan-Miklos, Shalina
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TITLE OF INVENTION: TREATMENT OF GASTROINTESTINAL DISORDERS
; FILE REFERENCE: 14184-043001
; CURRENT APPLICATION NUMBER: US/10/796.719
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US 10/766,735
; PRIOR FILING DATE: 2004-01-28
; PRIOR APPLICATION NUMBER: US 60/443,098
; PRIOR FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 60/471,288
; PRIOR FILING DATE: 2003-05-15
; PRIOR APPLICATION NUMBER: US 60/519,460
; PRIOR FILING DATE: 2003-11-12
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated peptide
US-10-796-719-27

Query Match      78.9%; Score 15; DB 17; Length 19;
```

```
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPAC 15
   |||||
Db 1 NSSNYCCCLCCNPAC 15
   |||||

RESULT 9
US-10-479-606-7
; Sequence 7, Application US/10479606
; Publication No. US20050032684A1
; GENERAL INFORMATION:
; APPLICANT: Savas, Yukeel
; APPLICANT: Mahajan-Miklos, Shalina
; TITLE OF INVENTION: Guanylate-cyclase C ligand, administered via the airways, for the
; FILE REFERENCE: 03100192aa
; CURRENT APPLICATION NUMBER: US/10/479,606
; CURRENT FILING DATE: 2003-12-04
; PRIOR APPLICATION NUMBER: DE10127119.0
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/DE02/02040
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-479-606-7

Query Match 78.9%; Score 15; DB 17; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPAC 15
   |||||
Db 1 NSSNYCCCLCCNPAC 15
   |||||

RESULT 10
US-10-796-719-39
; Sequence 39, Application US/10796719
; Publication No. US20050020811A1
; GENERAL INFORMATION:
; APPLICANT: Currie, Mark G.
; APPLICANT: Mahajan-Miklos, Shalina
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; FILE REFERENCE: 14184-043001
; CURRENT APPLICATION NUMBER: US/10/796,719
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US 10/766,735
; PRIOR FILING DATE: 2004-01-28
; PRIOR APPLICATION NUMBER: US 60/443,098
; PRIOR FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 60/471,288
; PRIOR FILING DATE: 2003-05-15
; PRIOR APPLICATION NUMBER: US 60/519,460
; PRIOR FILING DATE: 2003-11-12
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Synthetically generated peptide
US-10-796-719-39

Query Match 78.9%; Score 15; DB 17; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.6e-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPAC 15
   |||||
Db 1 NSSNYCCCLCCNPAC 15
   |||||

RESULT 11
US-10-796-719-40
; Sequence 40, Application US/10796719
; Publication No. US20050020811A1
; GENERAL INFORMATION:
; APPLICANT: Currie, Mark G.
; APPLICANT: Mahajan-Miklos, Shalina
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; FILE REFERENCE: 14184-043001
; CURRENT APPLICATION NUMBER: US/10/796,719
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US 10/766,735
; PRIOR FILING DATE: 2004-01-28
; PRIOR APPLICATION NUMBER: US 60/443,098
; PRIOR FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 60/471,288
; PRIOR FILING DATE: 2003-05-15
; PRIOR APPLICATION NUMBER: US 60/519,460
; PRIOR FILING DATE: 2003-11-12
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Synthetically generated peptide
US-10-796-719-40

Query Match 78.9%; Score 15; DB 17; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.6e-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSSNYCCCLCCNPAC 15
   |||||
Db 1 NSSNYCCCLCCNPAC 15
   |||||

RESULT 12
US-10-796-719-21
; Sequence 21, Application US/10796719
; Publication No. US20050020811A1
; GENERAL INFORMATION:
; APPLICANT: Currie, Mark G.
; APPLICANT: Mahajan-Miklos, Shalina
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; FILE REFERENCE: 14184-043001
; CURRENT APPLICATION NUMBER: US/10/796,719
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US 10/766,735
; PRIOR FILING DATE: 2004-01-28
; PRIOR APPLICATION NUMBER: US 60/443,098
; PRIOR FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 60/471,288
; PRIOR FILING DATE: 2003-05-15
; PRIOR APPLICATION NUMBER: US 60/519,460
; PRIOR FILING DATE: 2003-11-12
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-796-719-21
```

Query Match 78.9%; Score 15; DB 17; Length 72;  
Best Local Similarity 100.0%; Pred. No. 6.6e-09;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSSNYCCCLCCNPAC 15  
Db 54 NSSNYCCCLCCNPAC 68

## RESULT 13

US-10-621-684-28  
; Sequence 28, Application US/10621684  
; Publication No. US20040029182A1  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: ST Receptor Binding Compounds and  
; Methods of Using the Same

NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. US20040029182A1ris  
STREET: One Liberty Place, 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: WordPerfect 6.1

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/621,684  
FILING DATE: 17-Jul-2003  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/593,447A  
FILING DATE: 05-JAN-1996  
APPLICATION NUMBER: US 08/141,892  
FILING DATE: 26-OCT-1993

ATTORNEY/AGENT INFORMATION:  
NAME: DeLuca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-1702  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 28:

US-10-621-684-28

Query Match 73.7%; Score 14; DB 15; Length 17;  
Best Local Similarity 100.0%; Pred. No. 2.6e-08;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSNYCCCLCCNPAC 15  
Db 1 SSNYCCCLCCNPAC 14

## RESULT 14

US-10-621-684-33  
; Sequence 33, Application US/10621684  
; Publication No. US20040029182A1  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: ST Receptor Binding Compounds and  
; Methods of Using the Same

NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. US20040029182A1ris  
STREET: One Liberty Place, 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: WordPerfect 6.1

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/621,684  
FILING DATE: 17-Jul-2003  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/593,447A  
FILING DATE: 05-JAN-1996  
APPLICATION NUMBER: US 08/141,892  
FILING DATE: 26-OCT-1993

ATTORNEY/AGENT INFORMATION:  
NAME: DeLuca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-1702  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 33:  
US-10-621-684-33

Query Match 73.7%; Score 14; DB 15; Length 18;  
Best Local Similarity 100.0%; Pred. No. 2.7e-08;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSNYCCCLCCNPAC 15  
Db 1 SSNYCCCLCCNPAC 14

## RESULT 15

US-10-621-684-29  
; Sequence 29, Application US/10621684  
; Publication No. US20040029182A1  
; GENERAL INFORMATION:  
; APPLICANT: Waldman, Scott A.  
; TITLE OF INVENTION: ST Receptor Binding Compounds and  
; Methods of Using the Same

NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. US20040029182A1ris  
STREET: One Liberty Place, 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: WordPerfect 6.1

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/621,684  
FILING DATE: 17-Jul-2003  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/583,447A  
FILING DATE: 05-JAN-1996  
APPLICATION NUMBER: US 08/141,892  
FILING DATE: 26-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Deluca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-1702  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 29:  
US-10-621-684-29

Query Match 68.4%; Score 13; DB 15; Length 16;  
Best Local Similarity 100.0%; Pred. No. 2.9e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


Qy 3 SNYCCCLCCNPAC 15  
Db 1 SNYCCCLCCNPAC 13

Search completed: March 26, 2005, 17:44:31  
Job time : 61.4107 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2005, 17:10:31 ; Search time 19.6786 Seconds  
(without alignments)  
92.899 Million cell updates/sec

Title: US-10-775-481A-5  
Perfect score: 19  
Sequence: 1 NSSNYCCCLCCNPACNGCY 19  
Scoring table:  Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues  
Word size : 0  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Listing first 45 summaries

Database : PIR 79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 15    | 78.9        | 72     | 1 QHEC4  | heat-stable entero |
| 2          | 15    | 78.9        | 72     | 1 QHEC1B | heat-stable entero |
| 3          | 11    | 57.9        | 18     | 2 A60103 | heat-stable entero |
| 4          | 11    | 57.9        | 72     | 1 QHEC1  | heat-stable entero |
| 5          | 10    | 52.6        | 53     | 2 S68705 | heat-stable entero |
| 6          | 7     | 36.8        | 18     | 1 QHEC2  | heat-stable entero |
| 7          | 7     | 36.8        | 248    | 2 T19913 | hypothetical prote |
| 8          | 6     | 31.6        | 13     | 2 A28953 | alpha-conotoxin SI |
| 9          | 6     | 31.6        | 15     | 1 NTKNAG | alpha-conotoxin GI |
| 10         | 6     | 31.6        | 17     | 2 A54534 | heat-stable entero |
| 11         | 6     | 31.6        | 19     | 2 A44379 | alpha-conotoxin SI |
| 12         | 6     | 31.6        | 65     | 2 S34671 | heat-stable entero |
| 13         | 6     | 31.6        | 66     | 2 S31652 | enterotoxin - Vers |
| 14         | 6     | 31.6        | 71     | 2 S25659 | heat-stable entero |
| 15         | 6     | 31.6        | 78     | 1 QHVC1  | heat-stable entero |
| 16         | 6     | 31.6        | 124    | 1 NRPGB  | pancreatic ribonuc |
| 17         | 6     | 31.6        | 128    | 1 NRPGB  | pancreatic ribonuc |
| 18         | 6     | 31.6        | 136    | 2 T16269 | hypothetical prote |
| 19         | 6     | 31.6        | 408    | 2 T37929 | probable major fac |
| 20         | 6     | 31.6        | 795    | 2 T20091 | hypothetical prote |
| 21         | 5     | 26.3        | 75     | 2 H84631 | hypothetical prote |
| 22         | 5     | 26.3        | 90     | 2 D85845 | probable lysis pro |
| 23         | 5     | 26.3        | 110    | 2 I74319 | gene EHS-2 protein |
| 24         | 5     | 26.3        | 116    | 2 S77004 | umuC protein - Syn |
| 25         | 5     | 26.3        | 118    | 2 D97733 | hypothetical prote |
| 26         | 5     | 26.3        | 122    | 1 NRKGR  | pancreatic ribonuc |
| 27         | 5     | 26.3        | 124    | 1 NRBOB  | pancreatic ribonuc |
| 28         | 5     | 26.3        | 124    | 1 NRANE  | pancreatic ribonuc |
| 29         | 5     | 26.3        | 124    | 1 NRDEF  | pancreatic ribonuc |

|    |   |      |     |          |                    |
|----|---|------|-----|----------|--------------------|
| 30 | 5 | 26.3 | 124 | 1 NRDEN  | pancreatic ribonuc |
| 31 | 5 | 26.3 | 124 | 1 NRDEO  | pancreatic ribonuc |
| 32 | 5 | 26.3 | 124 | 1 NRDER  | pancreatic ribonuc |
| 33 | 5 | 26.3 | 124 | 1 NREKN  | pancreatic ribonuc |
| 34 | 5 | 26.3 | 124 | 1 NRGF   | pancreatic ribonuc |
| 35 | 5 | 26.3 | 124 | 1 NRGN   | pancreatic ribonuc |
| 36 | 5 | 26.3 | 124 | 1 NRGTH  | pancreatic ribonuc |
| 37 | 5 | 26.3 | 124 | 1 NRPBH  | pancreatic ribonuc |
| 38 | 5 | 26.3 | 124 | 1 NRSH   | pancreatic ribonuc |
| 39 | 5 | 26.3 | 124 | 1 NRWB   | pancreatic ribonuc |
| 40 | 5 | 26.3 | 124 | 2 S07141 | pancreatic ribonuc |
| 41 | 5 | 26.3 | 124 | 2 S08547 | pancreatic ribonuc |
| 42 | 5 | 26.3 | 124 | 2 S08546 | pancreatic ribonuc |
| 43 | 5 | 26.3 | 124 | 2 S08549 | pancreatic ribonuc |
| 44 | 5 | 26.3 | 124 | 2 C96762 | hypothetical prote |
| 45 | 5 | 26.3 | 125 | 2 A86872 | transcription regu |

ALIGNMENTS

RESULT 1

QHEC4

heat-stable enterotoxin STA4 precursor - Escherichia coli

C:Species: Escherichia coli

C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 09-Jul-2004

C:Accession: JT0373; A35978

R:Stieglitz, H.; Cervantes, L.; Robledo, R.; Fonseca, R.; Covarrubias, L.; Bolivar, F.;

Plasmid 20, 42-53, 1988

A>Title: Cloning, sequencing, and expression in ficoll-generated minicells of an Escheri

A:Reference number: JT0373; PMID:89202548; PMID:3071819

A:Accession: JT0373

A:Molecule type: DNA

A:Residues: 1-72 <STI>

A:Cross-references: UNIPROT:P07965; GB:J03311; MID:g147875; PIDN:AAA24652.1; PID:g147876

T:Zhou, X.; Shen, L.P.; Chi, C.W.

Toxicol 28, 453-456, 1990

A>Title: Isolation and nucleotide sequence determination of a gene encoding a heat-stabl

A:Reference number: A35978; PMID:90273381; PMID:2190361

A:Accession: A35978

A:Molecule type: DNA

A:Residues: 1-72 <ZHO>

C:Genetics:

A:Gene: estA4

C:Superfamily: heat-stable enterotoxin ST

C:Keywords: enterotoxin; heat-stable protein

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-53/Domain: propeptide #status predicted <PRO>

F:54-72/Product: heat-stable enterotoxin #status predicted <MAT>

F:59-64, 60-68, 63-71/Disulfide bonds: #status predicted

Query Match 78.9% Score 15; DB 1; Length 72;  
Best Local Similarity 100.0%; Pred. NO. 1.9e-10;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSSNYCCCLCCNPAC 15

Db 54 NSSNYCCCLCCNPAC 68

RESULT 2

QHEC1B

heat-stable enterotoxin ST-1b precursor - Escherichia coli

N:Alternate names: heat-stable enterotoxin ST-A2

C:Species: Escherichia coli

C>Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 09-Jul-2004

C:Accession: JS0292; A33068; A30567

R:Moseley, S.L.; Hardy, J.W.; Huq, M.I.; Echeverria, P.; Falkow, S.

Infect. Immun. 39, 1167-1174, 1983

A>Title: Isolation and nucleotide determination of a gene encoding a heat-stable enteroto

A:Reference number: JS0292; PMID:83184648; PMID:6341230

A:Accession: JS0292

A:Molecule type: DNA

A;Residues: 1-72 <WOS>  
 C;Cross-references: UNIPROT:Q47185; UNIPROT:P07965; GB:M34916; NID:G146407; PIDN:AAA23099  
 R;Dwarakanath, P.; Visweswariah, S.S.; Subrahmanyam, Y.V.B.K.; Shanthi, G.; Jagannatha, Gene 81, 219-226, 1989  
 A;Title: Cloning and hyperepression of a gene encoding the heat-stable toxin of *Escherichia coli*  
 A;Reference number: A33068; MUID:90034194; PMID:2680769  
 A;Accession: A33068  
 A;Molecule type: DNA  
 A;Residues: 1-18, 'A', 20-72 <DWA>  
 A;Cross-references: GB:M29255; NID:G148029; PIDN:AAA24686.1; PID:G148030  
 A;Note: the authors translated the codon AAG for residue 2 as Val and CTA for residue 34  
 R;Almota, S.; Takao, T.; Shimonishi, Y.; Hara, S.; Takeda, T.; Takeda, Y.; Miwatani, T. Eur. J. Biochem. 129, 257-263, 1982  
 A;Title: Amino acid sequence of heat-stable enterotoxin produced by human enterotoxigenic *Escherichia coli*  
 A;Reference number: A33067; MUID:83105138; PMID:6759126  
 A;Accession: A33067  
 A;Molecule type: protein  
 A;Residues: 54-72 <AIM>  
 R;Guzman-Verduzco, L.M.; Kuperstoch, Y.M. Infect. Immun. 57, 645-648, 1989  
 A;Title: Rectification of two *Escherichia coli* heat-stable enterotoxin allele sequences  
 A;Reference number: A30567; MUID:89108616; PMID:2643580  
 A;Accession: A30567  
 A;Molecule type: DNA  
 A;Residues: 1-18, 'A', 20-24, 'AG', 27-41, 'V', 43-44, 'N', 46, 'E', 48, 'S', 50-72 <GUZ>  
 A;Cross-references: GB:M18345; NID:G145862; PIDN:AAA23729.1; PID:G145863  
 C;Comment: This is one of the type I heat-stable enterotoxins that are methanol-soluble.  
 C;Genetics:  
 A;Gene: st  
 C;Superfamily: heat-stable enterotoxin ST  
 C;Keywords: enterotoxin; heat-stable protein  
 F;1-53/Domain: signal sequence and propeptide #status predicted <SIG>  
 F;54-72/Product: heat-stable enterotoxin ST-Ib #status experimental <MAT>  
 F;59-64,60-68,63-71/Disulfide bonds: #status experimental

Query Match 78.9%; Score 15; DB 1; Length 72;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-10;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSSNYCCELCCNPAC 15  
 |||||  
 DB 54 NSSNYCCELCCNPAC 68

RESULT 3  
 A60103  
 heat-stable enterotoxin ST-Ia - *Citrobacter freundii*  
 C;Species: *Citrobacter freundii*  
 C;Date: 10-Nov-1992 #sequence\_revision 10-Nov-1992 #text\_change 09-Jul-2004  
 C;Accession: A60103  
 R;Guarino, A.; Giannella, R.; Thompson, M.R. Infect. Immun. 57, 649-652, 1989  
 A;Title: *Citrobacter freundii* produces an 18-amino-acid heat-stable enterotoxin identical to *Escherichia coli* heat-stable enterotoxin ST-I  
 A;Reference number: A60103; MUID:89108617; PMID:2912902  
 A;Accession: A60103  
 A;Molecule type: protein  
 A;Residues: 1-18 <GUA>  
 A;Cross-references: UNIPROT:Q7M0U3  
 C;Superfamily: heat-stable enterotoxin ST

Query Match 57.9%; Score 11; DB 2; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YCCELCCNPAC 15  
 |||||  
 DB 4 YCCELCCNPAC 14

RESULT 4  
 QHSC1  
 heat-stable enterotoxin ST-I precursor - *Escherichia coli*  
 N;Alternate names: heat-stable enterotoxin estA1

C;Species: *Escherichia coli*  
 C;Date: 31-Aug-1980 #sequence\_revision 31-Aug-1980 #text\_change 09-Jul-2004  
 C;Accession: A01822; A30985; A36732; JT0374; I51932  
 R;So, M.; McCarthy, B.J. Proc. Natl. Acad. Sci. U.S.A. 77, 4011-4015, 1980  
 A;Title: Nucleotide sequence of the bacterial transposon Tn1681 encoding a heat-stable (S) enterotoxin  
 A;Reference number: A01822; MUID:81054703; PMID:6254008  
 A;Accession: A01822  
 A;Molecule type: DNA  
 A;Residues: 1-72 <LAZ>  
 A;Cross-references: UNIPROT:P01559; GB:V00612; GB:J01831; NID:G43704; PIDN:CAA23883.1; PID:G43704  
 R;Lazure, C.; Seidah, N.G.; Chretien, M.; Lallier, R.; St-Pierre, S. Can. J. Biochem. Cell Biol. 61, 287-292, 1983  
 A;Title: Primary structure determination of *Escherichia coli* heat-stable enterotoxin of *Escherichia coli*  
 A;Reference number: A30985; MUID:83284515; PMID:6349752  
 A;Accession: A30985  
 A;Molecule type: protein  
 A;Residues: 55-72 <LAZ2>  
 A;Experimental source: strain F11  
 R;Dallas, W.S. J. Bacteriol. 172, 5490-5493, 1990  
 A;Title: The heat-stable toxin I gene from *Escherichia coli* 18D.  
 A;Reference number: A36732; MUID:90368614; PMID:2203756  
 A;Accession: A36732  
 A;Molecule type: DNA  
 A;Residues: 1-72 <DAL>  
 A;Cross-references: GB:M58746; NID:G145860; PIDN:AAA62776.1; PID:G145861  
 A;Experimental source: strain 18D  
 R;Stieglitz, H.; Cervantes, L.; Robledo, R.; Fonseca, R.; Covarrubias, L.; Bolivar, F.; Plasmid 20, 42-53, 1988  
 A;Title: Cloning, sequencing, and expression in *Escherichia coli* heat-stable enterotoxin I  
 A;Reference number: JT0373; MUID:89202548; PMID:3071819  
 A;Accession: JT0374  
 A;Molecule type: DNA  
 A;Residues: 1-72 <STI>  
 R;Sekizaki, T.; Akashi, H.; Terakado, N. Am. J. Vet. Res. 46, 909-912, 1985  
 A;Title: Nucleotide sequences of the genes for *Escherichia coli* heat-stable enterotoxin I  
 A;Reference number: I51932; MUID:85249571; PMID:2990268  
 A;Accession: I51932  
 A;Status: translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-69, 'P', 71-72 <RES>  
 A;Cross-references: GB:M25607; NID:G147877; PIDN:AAA24653.1; PID:G147878  
 C;Comment: Both heat-stable and heat-labile enterotoxins are produced by pathogenic strains of *Escherichia coli*.  
 C;Superfamily: heat-stable enterotoxin ST  
 C;Keywords: enterotoxin; heat-stable protein  
 F;1-19/Domain: signal sequence #status predicted <SIG>  
 F;20-54/Domain: propeptide #status predicted <PRO>  
 F;55-72/Product: heat-stable enterotoxin ST-I #status experimental <MAT>  
 F;59-64,60-68,63-71/Disulfide bonds: #status predicted

Query Match 57.9%; Score 11; DB 1; Length 72;  
 Best Local Similarity 100.0%; Pred. No. 7.2e-06;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YCCELCCNPAC 15  
 |||||  
 DB 58 YCCELCCNPAC 68

RESULT 5  
 S68705  
 heat-stable enterotoxin Y-Stc - *Yersinia enterocolitica*  
 C;Species: *Yersinia enterocolitica*  
 C;Date: 25-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 13-Mar-1998  
 C;Accession: S68705  
 R;Yoshino, K.; Takao, T.; Huang, X.; Murata, H.; Nakao, H.; Takeda, T.; Shimonishi, Y. FEBS Lett. 362, 319-322, 1995  
 A;Title: Characterization of a highly toxic, large molecular size heat-stable enterotoxin  
 A;Reference number: S68705; MUID:95246844; PMID:7729521  
 A;Accession: S68705



A:Molecule type: protein  
A:Residues: 1-53 <YOS>  
A:Experimental source: strain 86-11  
C:Superfamily: heat-stable enterotoxin ST  
C:Keywords: enterotoxin; heat-stable protein  
F:41-46,42-50,45-53/Disulfide bonds: #status predicted

Query Match 52.6%; Score 10; DB 2; Length 53;  
Best Local Similarity 100.0%; Pred. No. 8e-05;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 CCELCNPAC 15  
|||||  
Db 41 CCELCNPAC 50

RESULT 6  
QHEC2  
heat-stable enterotoxin ST-2 - Escherichia coli  
C:Species: Escherichia coli  
C:Date: 06-Jul-1982 #sequence\_revision 06-Jul-1982 #text\_change 09-Jul-2004  
C:Accession: A01823  
R:Chan, S.K.; Giannella, R.A.  
J. Biol. Chem. 256, 7744-7746, 1981  
A:Title: Amino acid sequence of heat-stable enterotoxin produced by Escherichia coli pat  
A:Reference number: A01823; MUID:81264141; PMID:7021541  
A:Accession: A01823  
A:Molecule type: protein  
A:Residues: 1-18 <CHA>  
A:CROSS-references: UNIPROT:P01560  
A:Experimental source: strain 18D, serotype 0.42:k86:H37  
C:Comment: This enterotoxin is one of several, of differing molecular sizes, produced by  
idues of the heat-stable enterotoxin ST-1.  
C:Superfamily: heat-stable enterotoxin ST  
C:Keywords: enterotoxin; heat-stable protein  
F:1-18/Product: heat-stable enterotoxin ST-2 #status experimental <MAT>  
F:5-10,6-14,9-17/Disulfide bonds: #status predicted

Query Match 36.8%; Score 7; DB 1; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.098;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 YCCBLCC 11  
|||||  
Db 4 YCCBLCC 10

RESULT 7  
T19913  
hypothetical protein C43F9.5 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T19913  
R:Mortimore, B.  
submitted to the EMBL Data Library, November 1996  
A:Reference number: Z19195  
A:Accession: T19913  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-248 <WIL>  
A:CROSS-references: UNIPROT:Q9XUG6; EMBL:282262; PIDN:CAB05152.1; GSPDB:GN000022; CESP:C4  
A:Experimental source: clone C43F9  
C:Genetics:  
A:Gene: CESP:C43F9.5  
A:Map position: 4  
A:Introns: 47/1; 78/3; 118/1; 199/3

Query Match 36.8%; Score 7; DB 2; Length 248;  
Best Local Similarity 100.0%; Pred. No. 0.68;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSNYCCE 8  
|||||

Db 181 SSNYCCE 187

RESULT 8  
A28953  
alpha-conotoxin SI - cone shell (Conus striatus)  
C:Species: Conus striatus (striated cone)  
C:Date: 30-Jun-1989 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
C:Accession: A28953  
R:Zafaralla, G.C.; Ramilo, C.; Gray, W.R.; Karlstrom, R.; Olivera, B.M.; Cruz, L.J.  
Biochemistry 27, 7102-7105, 1988  
A:Title: Phylogenetic specificity of cholinergic ligands: alpha-conotoxin SI.  
A:Reference number: A28953; MUID:89062448; PMID:3196703  
A:Accession: A28953  
A:Molecule type: protein  
A:Residues: 1-13 <ZAF>  
A:CROSS-references: UNIPROT:P15471  
A:Note: This sequence was confirmed by chemical synthesis  
C:Comment: This paralytic toxin from a fish-hunting cone snail inhibits the acetylcholin  
C:Superfamily: alpha-conotoxin  
C:Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; postsynaptic neurot  
F:2-7,3-13/Disulfide bonds: #status experimental  
F:13/Modified site: amidated carboxyl end (Cys) #status experimental

Query Match 31.6%; Score 6; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 CCNPAC 15  
|||||  
Db 2 CCNPAC 7

RESULT 9  
NTRNG  
alpha-conotoxin GIA [validated] - cone shell (Conus geographus)  
A:Alternate names: alpha-C7X-GIA  
N:Contains: alpha-conotoxin GI  
C:Species: Conus geographus (geography cone)  
C:Date: 24-Sep-1981 #sequence\_revision 24-Sep-1981 #text\_change 09-Jul-2004  
C:Accession: A01782  
R:Gray, W.R.; Luque, A.; Olivera, B.M.; Barrett, J.; Cruz, L.J.  
J. Biol. Chem. 256, 4734-4740, 1981  
A:Title: Peptide toxins from Conus geographus venom.  
A:Reference number: A92320; MUID:81191854; PMID:7014556  
A:Accession: A01782  
A:Molecule type: protein  
A:Residues: 1-15 <GRA>  
A:CROSS-references: UNIPROT:P01519  
R:Gray, W.R.; Rivier, J.E.; Galyean, R.; Cruz, L.J.; Olivera, B.M.  
J. Biol. Chem. 258, 12247-12251, 1983  
A:Title: Conotoxin MI. Disulfide bonding and conformational states.  
A:Reference number: A92396; MUID:84032400; PMID:6630187  
A:Contents: annotation; disulfide bonds  
R:Guddat, L.W.; Shan, L.; Martin, J.L.; Edmundson, A.B.; Gray, W.R.  
submitted to the Brookhaven Protein Data Bank, May 1996  
A:Reference number: A66253; PDB:1NOT  
A:Contents: annotation; X-ray crystallography, 1.2 angstroms, residues 1-13  
R:Guddat, L.W.; Martin, J.A.; Shan, L.; Edmundson, A.B.; Gray, W.R.  
Biochemistry 35, 11329-11335, 1996  
A:Title: Three-dimensional structure of the alpha-conotoxin GI at 1.2 angstroms resoluti  
A:Reference number: A58592; MUID:96378624; PMID:8784187  
A:Contents: annotation; X-ray crystallography, 1.2 angstroms  
R:Pardi, A.; Galdes, A.; Florance, J.; Manicote, D.  
Biochemistry 28, 5494-5501, 1989  
A:Title: Solution structures of alpha-conotoxin GI determined by two-dimensional NMR spe  
A:Reference number: A30629; MUID:89375269; PMID:2775719  
A:Contents: annotation; conformation by (1)H-NMR  
C:Comment: This alpha-conotoxin, as an acetylcholine receptor inhibitor, is a postsynapt  
C:Superfamily: alpha-conotoxin  
C:Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; blocked carboxyl en  
F:1-15/Product: conotoxin GIA #status experimental <GIA>  
F:1-13/Product: conotoxin GI #status experimental <GIC>

F;2-7,3-13/Disulfide bonds: #link GFA #status predicted  
 F;2-7,3-13/Disulfide bonds: #link GFC #status experimental  
 F;13/Modified site: amidated carboxyl end (Cys) (amide in mature form from following gly  
 F;15/Modified site: blocked carboxyl end (Lys) (probably amidated) #status experimental

Query Match 31.6%; Score 6; DB 1; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 CCNPAC 15  
 |||||  
 Db 2 CCNPAC 7

RESULT 10  
 A54534  
 heat-stable enterotoxin - Vibrio mimicus (fragment)  
 C;Species: Vibrio mimicus  
 C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 03-May-1996  
 C;Accession: A54534  
 R;Arita, M.; Honda, T.; Miwatani, T.; Takeda, T.; Takao, T.; Shimonishi, Y.  
 FEMS Microbiol. Lett. 79, 105-110, 1991  
 A;Title: Purification and characterization of a heat-stable enterotoxin of Vibrio mimicus  
 A;Reference number: A54534  
 A;Accession: A54534  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-17 <ARI>  
 C;Superfamily: heat-stable enterotoxin ST

Query Match 31.6%; Score 6; DB 2; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 1.3;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 CCNPAC 15  
 |||||  
 Db 7 CCNPAC 12

RESULT 11  
 A44379  
 alpha-conotoxin SII - cone shell (Conus striatus)  
 C;Species: Conus striatus (striated cone)  
 C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 16-Jul-1999  
 C;Accession: A44379  
 R;Ramilo, C.A.; Zafaralla, G.C.; Nadasdi, L.; Hammerland, L.G.; Yoshikami, D.; Gray, W.R.  
 Biochemistry 31, 9919-9926, 1992  
 A;Title: Novel alpha- and omega-conotoxins from Conus striatus venom.  
 A;Reference number: A44379; MUID:93003172; PMID:1390774  
 A;Accession: A44379  
 A;Molecule type: protein  
 A;Residues: 1-19 <RAM>

A;Cross-references: CAS:143294-31-9; PIDN:AAB23762.1; PID:G257934  
 A;Experimental source: venom  
 A;Note: sequence extracted from NCBI backbone (NCBIP:116000); structure confirmed by che  
 C;Comment: This peptide is an acetylcholine receptor blocker.  
 C;Superfamily: alpha-conotoxin  
 C;Keywords: acetylcholine receptor inhibitor; postsynaptic neurotoxin; venom  
 F;2-18,3-8,4-14/Disulfide bonds: #status predicted  
 F;19/Modified site: amidated carboxyl end (Ser) #status absent

Query Match 31.6%; Score 6; DB 2; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 1.4;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 CCNPAC 15  
 |||||  
 Db 3 CCNPAC 8

RESULT 12  
 S34671  
 heat-stable enterotoxin - Vibrio cholerae (fragment)

C;Species: Vibrio cholerae  
 C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004  
 C;Accession: S34671  
 R;Rossolini, G.M.; Lombardi, G.; Guglielmetti, P.  
 submitted to the EMBL Data Library, July 1993

A;Description: Detection of the Vibrio cholerae heat-stable enterotoxin gene by polymerase  
 A;Reference number: S34671  
 A;Accession: S34671  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-65 <ROS>  
 A;Cross-references: UNIPROT:Q56643; EMBL:X74108  
 C;Superfamily: heat-stable enterotoxin ST

Query Match 31.6%; Score 6; DB 2; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 3.5;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 CCNPAC 15  
 |||||  
 Db 59 CCNPAC 64

RESULT 13  
 S31652  
 enterotoxin - Yersinia kristensenii  
 C;Species: Yersinia kristensenii  
 C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004  
 C;Accession: S31652  
 R;Ibrahim, A.; Liesack, W.; Stackebrandt, E.  
 submitted to the EMBL Data Library, November 1992  
 A;Reference number: S31652  
 A;Accession: S31652  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-66 <IBR>  
 A;Cross-references: UNIPROT:P31518; EMBL:X69218; NID:G48617; PIDN:CAA49152.1; PID:G48618  
 C;Superfamily: heat-stable enterotoxin ST

Query Match 31.6%; Score 6; DB 2; Length 66;  
 Best Local Similarity 100.0%; Pred. No. 3.6;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 CCNPAC 15  
 |||||  
 Db 58 CCNPAC 63

RESULT 14  
 S25659  
 heat-stable enterotoxin yst precursor - Yersinia enterocolitica  
 C;Species: Yersinia enterocolitica  
 C;Date: 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 09-Jul-2004  
 C;Accession: S25659; A41474; A23114; S65849  
 R;Ibrahim, A.; Liesack, W.; Pike, S.; Stackebrandt, E.  
 FEMS Microbiol. Lett. 97, 63-66, 1992  
 A;Title: The polymerase chain reaction: an epidemiological tool to differentiate between  
 A;Reference number: S25659  
 A;Accession: S25659  
 A;Status: preliminary; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-71 <IBR>  
 A;Cross-references: UNIPROT:P07593; EMBL:X65999; NID:G48611; PIDN:CAA4801.1; PID:G48612  
 R;Delor, I.; Kaechenbeek, A.; Wauters, G.; Cornelis, G.R.  
 Infect. Immun. 58, 2983-2988, 1990  
 A;Title: Nucleotide sequence of yst, the Yersinia enterocolitica gene encoding the heat-st  
 A;Reference number: A41474; MUID:90354067; PMID:2201642  
 A;Accession: A41474  
 A;Status: preliminary; not compared with conceptual translation  
 A;Molecule type: DNA  
 A;Residues: 1-47, 'S', 49-71 <DEL>

A;Cross-references: GB:U09235; NID:G487394; PIDN:AAA18472.1; PID:G487395  
 R;Takao, T.; Tominaga, N.; Yoshimura, S.; Shimonishi, Y.; Hara, S.; Inoue, T.; Miyama, A.

Eur. J. Biochem. 152, 199-206, 1985  
 A:Title: Isolation, primary structure and synthesis of heat-stable enterotoxin produced  
 A:Reference number: A23114; MUID:86004705; PMID:4043080  
 A:Accession: A23114  
 A:Molecule type: protein  
 A:Residues: 54-71 <TAK>  
 R:MKulskis, A.V.; Delor, I.; Ha Thi, V.; Cornelis, G.R.  
 Mol. Microbiol. 14, 905-915, 1994  
 A:Title: Regulation of the Yersinia enterocolitica enterotoxin Yst gene. Influence of g  
 A:Reference number: S65849; MUID:95231297; PMID:7715452  
 A:Accession: S65849  
 A:Status: preliminary; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-43 <MIK>  
 A:Cross-references: EMBL:U09235  
 C:Genetics:  
 A:Gene: yst  
 C:Superfamily: heat-stable enterotoxin ST  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-41/Domain: propeptide #status predicted <PRO>  
 F:42-71/Product: heat-stable enterotoxin yst #status predicted <MAT>

Query Match 31.6%; Score 6; DB 2; Length 71;  
 Best Local Similarity 100.0%; Pred. No. 3.8; Mismatches 0; Indels 0; Gaps 0;  
 Matches 6; Conservative 0;

Qy 10 CCNPAC 15  
 Db 63 CCNPAC 68  
 |||||

## RESULT 15

QHVC1  
 heat-stable enterotoxin ST precursor - Vibrio cholerae  
 C:Species: Vibrio cholerae  
 C:Date: 17-Mar-1987 #sequence revision 30-Jun-1992 #text\_change 09-Jul-2004  
 C:Accession: A41469; A01824; S34464; S34466; S34465; S34463  
 R:Ogawa, A.; Kato, J.I.; Watanabe, H.; Nair, B.G.; Takeda, T.  
 Infect. Immun. 58, 3325-3329, 1990  
 A:Title: Cloning and nucleotide sequence of a heat-stable enterotoxin gene from Vibrio c  
 A:Reference number: A41469; MUID:90382953; PMID:2205577  
 A:Accession: A41469  
 A:Molecule type: DNA  
 A:Residues: 1-78 <OGA>  
 A:Cross-references: UNIPROT:P04429; GB:M85198; GB:M36061; NID:G155237; PIDN:AAA64889.1;  
 R:Takao, T.; Shimonishi, Y.; Kobayashi, M.; Nishimura, O.; Arita, M.; Takeda, T.; Honda,  
 FEBS Lett. 193, 250-254, 1985  
 A:Title: Amino acid sequence of heat-stable enterotoxin produced by Vibrio cholerae non-  
 A:Reference number: A01824; MUID:86056320; PMID:4065341  
 A:Accession: A01824  
 A:Molecule type: protein  
 A:Residues: 62-78 <TAK>  
 A:Experimental source: non-O:1 serovar  
 R:Uoshino, K.; Miyachi, M.; Takao, T.; Bag, P.K.; Xiaozhe, H.; Nair, G.B.; Takeda, T.; S  
 FEBS Lett. 326, 83-86, 1993  
 A:Title: Purification and sequence determination of heat-stable enterotoxin elaborated b  
 A:Reference number: S34463; MUID:93314823; PMID:8325391  
 A:Accession: S34464  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 51-78 <YOS>  
 A:Accession: S34465  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 60-78 <Y02>  
 A:Accession: S34463  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 62-78 <Y04>

C:Superfamily: heat-stable enterotoxin ST  
 C:Keywords: enterotoxin; heat-stable protein  
 F:1-18/Domain: signal sequence #status predicted <SIG>  
 F:19-61/Domain: propeptide #status predicted <PRO>  
 F:62-78/Product: heat-stable enterotoxin ST #status experimental <MAT>  
 F:64-69,65-73,68-76/Disulfide bonds: #status predicted

Query Match 31.6%; Score 6; DB 1; Length 78;  
 Best Local Similarity 100.0%; Pred. No. 4;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 CCNPAC 15  
 Db 68 CCNPAC 73  
 |||||

Search completed: March 26, 2005, 17:26:00  
 Job time : 20.6786 secs

**This Page Blank (uspto)**